GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: July 31, 2002, 08:15:21 ; Search time 3626.68 Seconds
(without alignments)
17293.180 Million cell updates/sec

Title: US-09-674-817A-1
Perfect score: 2997
Sequence: 1 ggtcggggccggcgccgcgc.....aaaaaaaaaaaaaaaaaaaa 2997

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2997	100.0	2997	6	AX010481	AX010481 Sequence
2	2997	100.0	2997	6	AX010482	AX010482 Sequence
3	2997	100.0	2997	8	TAE307689	AJ307689 Triticum
4	1684.8	56.2	2590	8	AF438328	AF438328 Triticum
5	1614.8	53.9	2490	8	TAE301647	AJ301647 Triticum
6	1611.8	53.8	2437	6	AX010486	AX010486 Sequence
7	1510.8	50.4	2584	8	AB074189	AB074189 Hordeum v
8	1171.6	39.1	1708	8	AF142589	AF142589 Hordeum v
9	1134.2	37.8	1517	8	AF438329	AF438329 Triticum
10	1101.2	36.7	2288	8	AB015615	AB015615 Oryza sat
11	1074.4	35.8	1706	6	AX031277	AX031277 Sequence
12	1000	33.4	2712	6	AR072143	AR072143 Sequence
13	1000	33.4	2712	8	ZMU18908	U18908 Zea mays Su
14	932.8	31.1	983	8	AF142590	AF142590 Triticum
15	673.2	22.5	2706	6	AX006091	AX006091 Sequence
16	613.2	20.5	2806	6	AX006100	AX006100 Sequence
17	575.4	19.2	2133	6	A93395	A93395 Sequence 1
18	575.4	19.2	2133	6	AR161657	AR161657 Sequence
19	543.8	18.1	998	8	AF142591	AF142591 Solanum t
20	533.4	17.8	9289	6	AX031278	AX031278 Sequence
21	407.4	13.6	1663	6	A98394	A98394 Sequence 14
22	391.6	13.1	3291	8	AF142588	AF142588 Hordeum v
23	208.8	7.0	122349	1	D90908	D90908 Synechocyst
24	189.8	6.3	11779	8	AF030882	AF030882 Zea mays
25	149.2	5.0	2634	6	AX006092	AX006092 Sequence
26	143	4.8	13608	1	AE001623	AE001623 Chlamydia
c 27	143	4.8	13822	1	AE002198	AE002198 Chlamydop
28	143	4.8	300650	1	AP002546	AP002546 Chlamydop
29	129.6	4.3	1389	6	AX006101	AX006101 Sequence
30	129	4.3	253050	1	AP000984	AP000984 Sulfolobu
c 31	124	4.1	10608	1	AE002298	AE002298 Chlamydia
32	120	4.0	2523	6	AR072144	AR072144 Sequence
c 33	119.2	4.0	11652	1	AE001278	AE001278 Chlamydia
34	119.2	4.0	14262	1	AE006815	AE006815 Sulfolobu
35	119.2	4.0	100389	1	SS100KBFR	Y08256 S.solfatari
36	117.2	3.9	137740	1	D90900	D90900 Synechocyst
37	114.6	3.8	7218	1	D83245	D83245 Sulfolobus
38	110.8	3.7	92631	8	AP004469	AP004469 Lotus jap
39	110	3.7	110531	8	AF002109	AF002109 Arabidops
c 40	107.6	3.6	45590	1	SCBAC1A6	AL589708 Streptomy
c 41	105.2	3.5	10348	1	AE001888	AE001888 Deinococc
42	105	3.5	12519	1	AE006089	AE006089 Pasteurel
c 43	105	3.5	23506	1	AE008863	AE008863 Salmonell
c 44	104	3.5	9184	1	TSO001206	AJ001206 Streptomy
45	104	3.5	31336	1	SC4G10	AL356932 Streptomy

ALIGNMENTS

RESULT 1
AX010481
LOCUS AX010481 2997 bp DNA linear PAT 06-SEP-2000
DEFINITION Sequence 1 from Patent WO9958690.
ACCESSION AX010481
VERSION AX010481.1 GI:9997324
KEYWORDS .
SOURCE bread wheat.
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooideae; Triticeae; Triticum.
REFERENCE 1 (bases 1 to 2997)
AUTHORS Luetticke,S., Abel,G., Loerz,H. and Genschel,U.
TITLE Nucleic acid molecules which code for enzymes derived from wheat and which are involved in the synthesis of starch
JOURNAL Patent: WO 9958690-A 1 18-NOV-1999;
LUETTICKE STEPHANIE (DE); ABEL GERNOT (DE); LOERZ HORST (DE);
GENSCHEL ULRICH (DE); HOECHST SCHERING AGREVO GMBH (DE)
FEATURES Location/Qualifiers
source 1. .2997
/organism="Triticum aestivum"
/cultivar="Florida"

```
                     /db_xref="taxon:4565"
     CDS             join(<3. .296,397. .1617,2145. .2924)
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC07514.1"
                     /db_xref="GI:9997325"
                     /translation="SGPAPRLRRWRPNATAGKGVGEVCAAVVEAATKVEDEGEEDEPV
                     AEDRYALGGACRVLAGMPAPLGATALAGGVNFAVYSGGATAAALCLFTPEDLKADRVT
                     EEVPLDPLMNRTGNVWHVFIEGELHNMLYGYRFDGTFAPHCGHYLDVSNVVVDPYAKA
                     VISRGEYGVPARGNNCWPQMAGMIPLPYSTFDWEGDLPLRYPQKDLVIYEMHLRGFTK
                     HDSSNVEHPGTFIGAVSKLDYLKELGVNCIELMPCHEFNELEYSTSSSKMNFWGYSTI
                     NFFSPMTRYTSGGIKNCGRDAINEFKTFVREAHKRGIEVILDVVFNHTAEGNENGPIL
                     SFKGVDNTTYYMLAPKGEFYNYSGCGNTFNCNHPVVRQFIVDCLRYWVTEMHVDGFRF
                     DLASIMTRGSSLWDPVNVYGAPIEGDMITTGTPLVTPPLIDMISNDPILGGVKLIAEA
                     WDAGGLYQVGQFPHWNVWSEWNGKYRDIVRQFIKGTDGFAGGFAECLCGSPHLYQAGG
                     RKPWHSINFVCAHDGFTLADLVTYNKKYNLPNGENNRDGENHNLSWNCGEEGEFARLS
                     VKRLRKRQMRNFFVCLMVSQGVPMFYMGDEYGHTKGGNNNTYCHDSYVNYFRWDKKEQ
                     YSELHRFCCLMTKFRKECEGLGLEDFPTAKRLQWHGHQPGKPDWSENSRFVAFSMKDE
                     RQGEIYVAFNTSHLPAVVELPERAGRRWEPVVDTGKPAPYDFLTDDLPDRALTIHQFS
                     HFLYSNLYPMLSYSSVILVLRPDV"
BASE COUNT      803 a    636 c    762 g    796 t
ORIGIN
```

```
  Query Match             100.0%;  Score 2997;  DB 6;  Length 2997;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 2997;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ggtcggggccggcgccgcgcctgcgacggtggcgacccaatgcgacggcggggaaggggg 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 GGTCGGGGCCGGCGCCGCGCCTGCGACGGTGGCGACCCAATGCGACGGCGGGGAAGGGGG 60

Qy      61 tcggcgaggtgtgcgccgcggttgtcgaggcggcgacgaaggtagaggacgaggggggagg 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 TCGGCGAGGTGTGCGCCGCGGTTGTCGAGGCGGCGACGAAGGTAGAGGACGAGGGGGAGG 120

Qy     121 aggacgagccggtggcggaggacaggtacgcgctcggcggcgcgtgcagggtgctcgccg 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 AGGACGAGCCGGTGGCGGAGGACAGGTACGCGCTCGGCGGCGCGTGCAGGGTGCTCGCCG 180

Qy     181 gaatgcccgcgccgctgggcgccaccgcgctcgccggcggggtcaatttcgccgtctatt 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 GAATGCCCGCGCCGCTGGGCGCCACCGCGCTCGCCGGCGGGGTCAATTTCGCCGTCTATT 240

Qy     241 ccggcggagccaccgccgcggcgctctgcctcttcacgccagaagatctcaaggcggtgg 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CCGGCGGAGCCACCGCCGCGGCGCTCTGCCTCTTCACGCCAGAAGATCTCAAGGCGGTGG 300

Qy     301 ggttgcctcccgagtagagttcatcagctttgcgtgcgccgcgcgccccttttttgggcc 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 GGTTGCCTCCCGAGTAGAGTTCATCAGCTTTGCGTGCGCCGCGCGCCCCTTTTTTGGGCC 360

Qy     361 tgcaatttaagttttgtactggggcaaatgctgcaggatagggtgaccgaggaggttccc 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 TGCAATTTAAGTTTTGTACTGGGGCAAATGCTGCAGGATAGGGTGACCGAGGAGGTTCCC 420

Qy     421 cttgacccccctgatgaatcggaccgggaacgtgtggcatgtcttcatcgaaggcgagctg 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 CTTGACCCCCTGATGAATCGGACCGGGAACGTGTGGCATGTCTTCATCGAAGGCGAGCTG 480

Qy     481 cacaacatgctttacgggtacaggttcgacggcacctttgctcctcactgcgggcactac 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 CACAACATGCTTTACGGGTACAGGTTCGACGGCACCTTTGCTCCTCACTGCGGGCACTAC 540

Qy     541 cttgatgtttccaatgtcgtggtggatccttatgctaaggcagtgataagccgaggggag 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 CTTGATGTTTCCAATGTCGTGGTGGATCCTTATGCTAAGGCAGTGATAAGCCGAGGGGAG 600

Qy     601 tatggtgttccagcgcgtggtaacaattgctggcctcagatggctggcatgatccctctt 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 TATGGTGTTCCAGCGCGTGGTAACAATTGCTGGCCTCAGATGGCTGGCATGATCCCTCTT 660

Qy     661 ccatatagcacgtttgattgggaaggcgacctacctctaagatatcctcaaaaggacctg 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 CCATATAGCACGTTTGATTGGGAAGGCGACCTACCTCTAAGATATCCTCAAAAGGACCTG 720

Qy     721 gtaatatatgagatgcacttgcgtggattcacgaagcatgattcaagcaatgtagaacat 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 GTAATATATGAGATGCACTTGCGTGGATTCACGAAGCATGATTCAAGCAATGTAGAACAT 780

Qy     781 ccgggtactttcattggagctgtgtcgaagcttgactatttgaaggagcttggagttaat 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 CCGGGTACTTTCATTGGAGCTGTGTCGAAGCTTGACTATTTGAAGGAGCTTGGAGTTAAT 840

Qy     841 tgtattgaattaatgccctgccatgagttcaacgagctggagtactcaacctcttcttcc 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 TGTATTGAATTAATGCCCTGCCATGAGTTCAACGAGCTGGAGTACTCAACCTCTTCTTCC 900

Qy     901 aagatgaacttttggggatattctaccataaacttcttttcaccaatgacaagatacaca 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 AAGATGAACTTTTGGGGATATTCTACCATAAACTTCTTTTCACCAATGACAAGATACACA 960

Qy     961 tcaggcgggataaaaaaactgtgggcgtgatgccataaatgagttcaaaacttttgtaaga 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 TCAGGCGGGATAAAAAACTGTGGGCGTGATGCCATAAATGAGTTCAAAACTTTTGTAAGA 1020

Qy    1021 gaggctcacaaacggggaattgaggtgatcctggatgttgtcttcaaccatacagctgag 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 GAGGCTCACAAACGGGGAATTGAGGTGATCCTGGATGTTGTCTTCAACCATACAGCTGAG 1080

Qy    1081 ggtaatgagaatggtccaatattatcatttaaggggtcgataatactacatactatatg 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 GGTAATGAGAATGGTCCAATATTATCATTTAAGGGGTCGATAATACTACATACTATATG 1140

Qy    1141 cttgcacccaagggagagttttataactattctggctgtgggaataccttcaactgtaat 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 CTTGCACCCAAGGGAGAGTTTTATAACTATTCTGGCTGTGGGAATACCTTCAACTGTAAT 1200

Qy    1201 catcctgtggttcgtcaattcattgtagattgtttaagatactgggtgacggaaatgcat 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 CATCCTGTGGTTCGTCAATTCATTGTAGATTGTTTAAGATACTGGGTGACGGAAATGCAT 1260

Qy    1261 gttgatggttttcgttttgatcttgcatccataatgaccagaggttccagtctgtgggat 1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 GTTGATGGTTTTCGTTTTGATCTTGCATCCATAATGACCAGAGGTTCCAGTCTGTGGGAT 1320

Qy    1321 ccagttaacgtgtatggagctccaatagaaggtgacatgatcacaacagggacacctctt 1380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321 CCAGTTAACGTGTATGGAGCTCCAATAGAAGGTGACATGATCACAACAGGGACACCTCTT 1380

Qy    1381 gttactccaccacttattgacatgatcagcaatgacccaattcttggaggcgtcaagctc 1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1381 GTTACTCCACCACTTATTGACATGATCAGCAATGACCCAATTCTTGGAGGCGTCAAGCTC 1440

Qy    1441 attgctgaagcatgggatgcaggaggcctctatcaagtaggtcaattccctcactggaat 1500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1441 ATTGCTGAAGCATGGGATGCAGGAGGCCTCTATCAAGTAGGTCAATTCCCTCACTGGAAT 1500

Qy    1501 gtttggtctgagtggaatgggaagtaccgggacattgtgcgtcaattcattaaaggcact 1560
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1501 GTTTGGTCTGAGTGGAATGGGAAGTACCGGGACATTGTGCGTCAATTCATTAAAGGCACT 1560

Qy    1561 gatggatttgctggtggttttgccgaatgtctttgtggaagtccacacctataccaggta 1620
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1561 GATGGATTTGCTGGTGGTTTTGCCGAATGTCTTTGTGGAAGTCCACACCTATACCAGGTA 1620

Qy    1621 agttgtggcaatacttgtaaatgagttgagtgaatgtcacctggatttttttatatatacc 1680
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1621 AGTTGTGGCAATACTTGTAAATGAGTTGAGTGAATGTCACCTGGATTTTTTATATATACC 1680

Qy    1681 acatgatgatacacatctaaatatataacaatcatagtgtatgcatatgcatttggctaa 1740
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1681 ACATGATGATACACATCTAAATATATAACAATCATAGTGTATGCATATGCATTTGGCTAA 1740

Qy    1741 gaagtattagtgtatacactagtgctatatataggttttaacacccaacttgccaatgaa 1800
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   1741 GAAGTATTAGTGTATACACTAGTGCTATATATAGGTTTTAACACCCAACTTGCCAATGAA 1800

Qy   1801 ggaacatagggctttctagttatcttatttatttgtccggtgaataatccactgaaaaat 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 GGAACATAGGGCTTTCTAGTTATCTTATTTATTTGTCCGGTGAATAATCCACTGAAAAAT 1860

Qy   1861 tccagccatgtcatttttttaggggggagaagaaactatattgatttgccccccctaaaag 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 TCCAGCCATGTCATTTTTTTAGGGGGGGAGAAGAAACTATATTGATTTGCCCCCCTAAAAG 1920

Qy   1921 aagccatctcagaattcataggtaagttgctttttctgtaaagaaaggaaaacgacttcat 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 AAGCCATCTCAGAATTCATAGGTAAGTTGCTTTTCTGTAAAGAAAGGAAAACGACTTCAT 1980

Qy   1981 actttctatcggtgctaacttagctcgatgtatatttgtaagatgaatgccaaatttaat 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 ACTTTCTATCGGTGCTAACTTAGCTCGATGTATATTTGTAAGATGAATGCCAAATTTAAT 2040

Qy   2041 ttgtcggataatttgatctgttattcacaaatttctatttggtttctctagaaatcaaac 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 TTGTCGGATAATTTGATCTGTTATTCACAAATTTCTATTTGGTTTCTCTAGAAATCAAAC 2100

Qy   2101 cagtaacttgttattggcactgcaacttcttattgattaatcaggcaggaggaaggaaac 2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 CAGTAACTTGTTATTGGCACTGCAACTTCTTATTGATTAATCAGGCAGGAGGAAGGAAAC 2160

Qy   2161 cttggcacagtatcaactttgtatgtgcacatgatggatttacactggctgatttggtaa 2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 CTTGGCACAGTATCAACTTTGTATGTGCACATGATGGATTTACACTGGCTGATTTGGTAA 2220

Qy   2221 catataataagaagtacaatttaccaaatggggagaacaacagagatggagaaaatcaca 2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 CATATAATAAGAAGTACAATTTACCAAATGGGGAGAACAACAGAGATGGAGAAAATCACA 2280

Qy   2281 atcttagctggaattgtggggaggaaggagaattcgcaagattgtctgtcaaaagattga 2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 ATCTTAGCTGGAATTGTGGGGAGGAAGGAGAATTCGCAAGATTGTCTGTCAAAAGATTGA 2340

Qy   2341 ggaagaggcagatgcgcaatttctttgtttgtctcatggtttctcaaggagttccaatgt 2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 GGAAGAGGCAGATGCGCAATTTCTTTGTTTGTCTCATGGTTTCTCAAGGAGTTCCAATGT 2400

Qy   2401 tctacatgggtgatgaatatggccacacaaaaggggcaacaacaatacatactgccatg 2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401 TCTACATGGGTGATGAATATGGCCACACAAAAGGGGCAACAACAATACATACTGCCATG 2460

Qy   2461 attcttatgtcaattattttcgctgggataaaaaaagaacaatactctgagttgcaccgat 2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2461 ATTCTTATGTCAATTATTTTCGCTGGGATAAAAAAGAACAATACTCTGAGTTGCACCGAT 2520

Qy   2521 tctgctgcctcatgaccaaattccgcaaggagtgcgagggtcttggccttgaggactttc 2580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2521 TCTGCTGCCTCATGACCAAATTCCGCAAGGAGTGCGAGGGTCTTGGCCTTGAGGACTTTC 2580

Qy   2581 caacggccaaacggctgcagtggcatggtcatcagcctgggaagcctgattggtctgaga 2640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2581 CAACGGCCAAACGGCTGCAGTGGCATGGTCATCAGCCTGGGAAGCCTGATTGGTCTGAGA 2640

Qy   2641 atagccgattcgttgccttttccatgaaagatgaaagacagggcgagatctatgtggcct 2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2641 ATAGCCGATTCGTTGCCTTTTCCATGAAAGATGAAAGACAGGGCGAGATCTATGTGGCCT 2700

Qy   2701 tcaacaccagccacttaccggccgttgttgagctcccagagcgcgcagggcgccggtggg 2760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2701 TCAACACCAGCCACTTACCGGCCGTTGTTGAGCTCCCAGAGCGCGCAGGGCGCCGGTGGG 2760

Qy   2761 aaccggtggtggacacaggcaagccagcaccatacgacttcctcaccgacgacttacctg 2820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2761 AACCGGTGGTGGACACAGGCAAGCCAGCACCATACGACTTCCTCACCGACGACTTACCTG 2820

Qy   2821 atcgcgctctcaccatacaccagttctcgcatttcctctactccaacctctaccccatgc 2880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2821 ATCGCGCTCTCACCATACACCAGTTCTCGCATTTCCTCTACTCCAACCTCTACCCCATGC 2880
```

```
Qy   2881 tcagctactcatcggtcatcctagtattgcgccctgatgtttgagagaccaatatatataca 2940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2881 TCAGCTACTCATCGGTCATCCTAGTATTGCGCCCTGATGTTTGAGAGACCAATATATATACA 2940

Qy   2941 gtaaataatatgtctatatgtaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 2997
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2941 GTAAATAATATGTCTATATGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2997


RESULT   2
AX010482
LOCUS       AX010482               2997 bp    DNA     linear   PAT 06-SEP-2000
DEFINITION  Sequence 2 from Patent WO9958690.
ACCESSION   AX010482
VERSION     AX010482.1  GI:9997326
KEYWORDS    .
SOURCE      bread wheat.
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Triticum.
REFERENCE   1  (bases 1 to 2997)
  AUTHORS   Luetticke,S., Abel,G., Loerz,H. and Genschel,U.
  TITLE     Nucleic acid molecules which code for enzymes derived from wheat
            and which are involved in the synthesis of starch
  JOURNAL   Patent: WO 9958690-A 2 18-NOV-1999;
            LUETTICKE STEPHANIE (DE); ABEL GERNOT (DE); LOERZ HORST (DE);
            GENSCHEL ULRICH (DE); HOECHST SCHERING AGREVO GMBH (DE)
FEATURES             Location/Qualifiers
     source          1. .2997
                     /organism="Triticum aestivum"
                     /cultivar="Florida"
                     /db_xref="taxon:4565"
BASE COUNT      803 a    636 c    762 g    796 t
ORIGIN


  Query Match           100.0%;  Score 2997;  DB 6;  Length 2997;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 2997;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ggtcggggccggcgccgcgcctgcgacggtggcgacccaatgcgacggcggggaaggggg 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 GGTCGGGGCCGGCGCCGCGCCTGCGACGGTGGCGACCCAATGCGACGGCGGGGAAGGGGG 60

Qy      61 tcggcgaggtgtgcgccgcggttgtcgaggcggcgacgaaggtagaggacgaggggggagg 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 TCGGCGAGGTGTGCGCCGCGGTTGTCGAGGCGGCGACGAAGGTAGAGGACGAGGGGGAGG 120

Qy     121 aggacgagccggtggcggaggacaggtacgcgctcggcggcgcgtgcagggtgctcgccg 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 AGGACGAGCCGGTGGCGGAGGACAGGTACGCGCTCGGCGGCGCGTGCAGGGTGCTCGCCG 180

Qy     181 gaatgcccgcgccgctgggcgccaccgcgctcgccggcggggtcaatttcgccgtctatt 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 GAATGCCCGCGCCGCTGGGCGCCACCGCGCTCGCCGGCGGGGTCAATTTCGCCGTCTATT 240

Qy     241 ccggcggagccaccgccgcggcgctctgcctcttcacgccagaagatctcaaggcggtgg 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CCGGCGGAGCCACCGCCGCGGCGCTCTGCCTCTTCACGCCAGAAGATCTCAAGGCGGTGG 300

Qy     301 ggttgcctcccgagtagagttcatcagctttgcgtgcgccgcgcgccccttttttgggcc 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 GGTTGCCTCCCGAGTAGAGTTCATCAGCTTTGCGTGCGCCGCGCGCCCCTTTTTTGGGCC 360

Qy     361 tgcaatttaagttttgtactggggcaaatgctgcaggatagggtgaccgaggaggttccc 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 TGCAATTTAAGTTTTGTACTGGGGCAAATGCTGCAGGATAGGGTGACCGAGGAGGTTCCC 420

Qy     421 cttgacccccctgatgaatcggaccgggaacgtgtggcatgtcttcatcgaaggcgagctg 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 CTTGACCCCCTGATGAATCGGACCGGGAACGTGTGGCATGTCTTCATCGAAGGCGAGCTG 480
```

```
Qy    481 cacaacatgctttacgggtacaggttcgacggcacctttgctcctcactgcgggcactac 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CACAACATGCTTTACGGGTACAGGTTCGACGGCACCTTTGCTCCTCACTGCGGGCACTAC 540

Qy    541 cttgatgtttccaatgtcgtggtggatccttatgctaaggcagtgataagccgaggggag 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CTTGATGTTTCCAATGTCGTGGTGGATCCTTATGCTAAGGCAGTGATAAGCCGAGGGGAG 600

Qy    601 tatggtgttccagcgcgtggtaacaattgctggcctcagatggctggcatgatccctctt 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 TATGGTGTTCCAGCGCGTGGTAACAATTGCTGGCCTCAGATGGCTGGCATGATCCCTCTT 660

Qy    661 ccatatagcacgtttgattgggaaggcgacctacctctaagatatcctcaaaaggacctg 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 CCATATAGCACGTTTGATTGGGAAGGCGACCTACCTCTAAGATATCCTCAAAAGGACCTG 720

Qy    721 gtaatatatgagatgcacttgcgtggattcacgaagcatgattcaagcaatgtagaacat 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GTAATATATGAGATGCACTTGCGTGGATTCACGAAGCATGATTCAAGCAATGTAGAACAT 780

Qy    781 ccgggtactttcattggagctgtgtcgaagcttgactatttgaaggagcttggagttaat 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 CCGGGTACTTTCATTGGAGCTGTGTCGAAGCTTGACTATTTGAAGGAGCTTGGAGTTAAT 840

Qy    841 tgtattgaattaatgccctgccatgagttcaacgagctggagtactcaacctcttcttcc 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 TGTATTGAATTAATGCCCTGCCATGAGTTCAACGAGCTGGAGTACTCAACCTCTTCTTCC 900

Qy    901 aagatgaacttttggggatattctaccataaacttcttttcaccaatgacaagatacaca 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 AAGATGAACTTTTGGGGATATTCTACCATAAACTTCTTTTCACCAATGACAAGATACACA 960

Qy    961 tcaggcgggataaaaaaactgtgggcgtgatgccataaatgagttcaaaacttttgtaaga 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 TCAGGCGGGATAAAAAACTGTGGGCGTGATGCCATAAATGAGTTCAAAACTTTTGTAAGA 1020

Qy   1021 gaggctcacaaacggggaattgaggtgatcctggatgttgtcttcaaccatacagctgag 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 GAGGCTCACAAACGGGGAATTGAGGTGATCCTGGATGTTGTCTTCAACCATACAGCTGAG 1080

Qy   1081 ggtaatgagaatggtccaatattatcatttaaggggtcgataatactacatactatatg 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 GGTAATGAGAATGGTCCAATATTATCATTTAAGGGGGTCGATAATACTACATACTATATG 1140

Qy   1141 cttgcacccaagggagagttttataactattctggctgtgggaatacttcaactgtaat 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 CTTGCACCCAAGGGAGAGTTTTATAACTATTCTGGCTGTGGGAATACCTTCAACTGTAAT 1200

Qy   1201 catcctgtggttcgtcaattcattgtagattgtttaagatactgggtgacggaaatgcat 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 CATCCTGTGGTTCGTCAATTCATTGTAGATTGTTTAAGATACTGGGTGACGGAAATGCAT 1260

Qy   1261 gttgatggttttcgttttgatcttgcatccataatgaccagaggttccagtctgtgggat 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 GTTGATGGTTTTCGTTTTGATCTTGCATCCATAATGACCAGAGGTTCCAGTCTGTGGGAT 1320

Qy   1321 ccagttaacgtgtatggagctccaatagaaggtgacatgatcacaacagggacacctctt 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 CCAGTTAACGTGTATGGAGCTCCAATAGAAGGTGACATGATCACAACAGGGACACCTCTT 1380

Qy   1381 gttactccaccacttattgacatgatcagcaatgacccaattcttggaggcgtcaagctc 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 GTTACTCCACCACTTATTGACATGATCAGCAATGACCCAATTCTTGGAGGCGTCAAGCTC 1440

Qy   1441 attgctgaagcatgggatgcaggaggcctctatcaagtaggtcaattccctcactggaat 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 ATTGCTGAAGCATGGGATGCAGGAGGCCTCTATCAAGTAGGTCAATTCCCTCACTGGAAT 1500

Qy   1501 gtttggtctgagtggaatgggaagtaccgggacattgtgcgtcaattcattaaaggcact 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 GTTTGGTCTGAGTGGAATGGGAAGTACCGGGACATTGTGCGTCAATTCATTAAAGGCACT 1560

Qy   1561 gatggatttgctggtggttttgccgaatgtctttgtggaagtccacacctataccaggta 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 GATGGATTTGCTGGTGGTTTTGCCGAATGTCTTTGTGGAAGTCCACACCTATACCAGGTA 1620

Qy   1621 agttgtggcaatacttgtaaatgagttgagtgaatgtcacctggatttttttatatatacc 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 AGTTGTGGCAATACTTGTAAATGAGTTGAGTGAATGTCACCTGGATTTTTTATATATACC 1680

Qy   1681 acatgatgatacacatctaaatatataacaatcatagtgtatgcatatgcatttggctaa 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 ACATGATGATACACATCTAAATATATAACAATCATAGTGTATGCATATGCATTTGGCTAA 1740

Qy   1741 gaagtattagtgtatacactagtgctatatataggttttaacacccaacttgccaatgaa 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 GAAGTATTAGTGTATACACTAGTGCTATATATAGGTTTTAACACCCAACTTGCCAATGAA 1800

Qy   1801 ggaacatagggctttctagttatcttatttatttgtccggtgaataatccactgaaaaat 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 GGAACATAGGGCTTTCTAGTTATCTTATTTATTTGTCCGGTGAATAATCCACTGAAAAAT 1860

Qy   1861 tccagccatgtcatttttttaggggggggagaagaaactatattgatttgccccccctaaaag 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 TCCAGCCATGTCATTTTTTAGGGGGGGAGAAGAAACTATATTGATTTGCCCCCCCTAAAAG 1920

Qy   1921 aagccatctcagaattcataggtaagttgctttttctgtaaagaaaggaaaacgacttcat 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 AAGCCATCTCAGAATTCATAGGTAAGTTGCTTTTCTGTAAAGAAAGGAAAACGACTTCAT 1980

Qy   1981 actttctatcggtgctaacttagctcgatgtatatttgtaagatgaatgccaaatttaat 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 ACTTTCTATCGGTGCTAACTTAGCTCGATGTATATTTGTAAGATGAATGCCAAATTTAAT 2040

Qy   2041 ttgtcggataatttgatctgttattcacaaatttctatttggtttctctagaaatcaaac 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 TTGTCGGATAATTTGATCTGTTATTCACAAATTTCTATTTGGTTTCTCTAGAAATCAAAC 2100

Qy   2101 cagtaacttgttattggcactgcaacttcttattgattaatcaggcaggaggaaggaaac 2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 CAGTAACTTGTTATTGGCACTGCAACTTCTTATTGATTAATCAGGCAGGAGGAAGGAAAC 2160

Qy   2161 cttggcacagtatcaactttgtatgtgcacatgatggatttacactggctgatttggtaa 2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 CTTGGCACAGTATCAACTTTGTATGTGCACATGATGGATTTACACTGGCTGATTTGGTAA 2220

Qy   2221 catataataagaagtacaatttaccaaatggggagaacaacagagatggagaaaatcaca 2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 CATATAATAAGAAGTACAATTTACCAAATGGGGAGAACAACAGAGATGGAGAAAATCACA 2280

Qy   2281 atcttagctggaattgtggggaggaaggagaattcgcaagattgtctgtcaaaagattga 2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 ATCTTAGCTGGAATTGTGGGGAGGAAGGAGAATTCGCAAGATTGTCTGTCAAAAGATTGA 2340

Qy   2341 ggaagaggcagatgcgcaatttctttgtttgtctcatggtttctcaaggagttccaatgt 2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 GGAAGAGGCAGATGCGCAATTTCTTTGTTTGTCTCATGGTTTCTCAAGGAGTTCCAATGT 2400

Qy   2401 tctacatgggtgatgaatatggccacacaaaaggggcaacaacaatacatactgccatg 2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401 TCTACATGGGTGATGAATATGGCCACACAAAAGGGGCAACAACAATACATACTGCCATG 2460

Qy   2461 attcttatgtcaattatttttcgctgggataaaaaagaacaatactctgagttgcaccgat 2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2461 ATTCTTATGTCAATTATTTTCGCTGGGATAAAAAAGAACAATACTCTGAGTTGCACCGAT 2520

Qy   2521 tctgctgcctcatgaccaaattccgcaaggagtgcgagggtcttggccttgaggactttc 2580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2521 TCTGCTGCCTCATGACCAAATTCCGCAAGGAGTGCGAGGGTCTTGGCCTTGAGGACTTTC 2580

Qy   2581 caacggccaaacggctgcagtggcatggtcatcagcctgggaagcctgattggtctgaga 2640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2581 CAACGGCCAAACGGCTGCAGTGGCATGGTCATCAGCCTGGGAAGCCTGATTGGTCTGAGA 2640

Qy   2641 atagccgattcgttgccttttccatgaaagatgaaagacagggcgagatctatgtggcct 2700
```

```
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2641 ATAGCCGATTCGTTGCCTTTTCCATGAAAGATGAAAGACAGGGCGAGATCTATGTGGCCT 2700

Qy    2701 tcaacaccagccacttaccggccgttgttgagctcccagagcgcgcagggcgccggtggg 2760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2701 TCAACACCAGCCACTTACCGGCCGTTGTTGAGCTCCCAGAGCGCGCAGGGCGCCGGTGGG 2760

Qy    2761 aaccggtggtggacacaggcaagccagcaccatacgacttcctcaccgacgacttacctg 2820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2761 AACCGGTGGTGGACACAGGCAAGCCAGCACCATACGACTTCCTCACCGACGACTTACCTG 2820

Qy    2821 atcgcgctctcaccatacaccagttctcgcatttcctctactccaacctctaccccatgc 2880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2821 ATCGCGCTCTCACCATACACCAGTTCTCGCATTTCCTCTACTCCAACCTCTACCCCATGC 2880

Qy    2881 tcagctactcatcggtcatcctagtattgcgccctgatgtttgagagaccaatatatataca 2940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2881 TCAGCTACTCATCGGTCATCCTAGTATTGCGCCCTGATGTTTGAGAGACCAATATATATACA 2940

Qy    2941 gtaaataatatgtctatatgtaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 2997
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2941 GTAAATAATATGTCTATATGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2997


RESULT   3
TAE307689
LOCUS       TAE307689               2997 bp    mRNA    linear   PLN 17-DEC-2001
DEFINITION  Triticum aestivum partial precursor RNA for isoamylase (iso-1b
            gene).
ACCESSION   AJ307689
VERSION     AJ307689.1  GI:14331017
KEYWORDS    iso-1b gene; isoamylase.
SOURCE      bread wheat.
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Triticum.
REFERENCE   1  (sites)
  AUTHORS   Genschel,U., Abel,G., Loerz,H. and Luetticke,S.
  TITLE     The sugary-type Isoamylases in wheat
  JOURNAL   Unpublished
REFERENCE   2  (sites)
  AUTHORS   Abel,G., Loerz,H. and Luetticke,S.
  TITLE     The Sugary-type Isoamylase in Wheat
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 2997)
  AUTHORS   Luetticke,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-JUN-2001) Luetticke S., Insitut fuer Allgemeine
            Botanik, Universitaet Hamburg, Ohnhorststrasse 18, Hamburg, 20609,
            GERMANY
COMMENT     Related entry AJ301647.
FEATURES             Location/Qualifiers
     source          1. .2997
                     /organism="Triticum aestivum"
                     /variety="Florida"
                     /db_xref="taxon:4565"
                     /tissue_type="developing kernel"
     precursor_RNA   <1. .2924
                     /gene="iso-1b"
     exon            <1. .297
                     /gene="iso-1b"
                     /number=1
     mRNA            join(<1. .297,398. .1618,2146. .2924)
                     /gene="iso-1b"
     gene            1. .2924
                     /gene="iso-1b"
     CDS             join(<1. .297,398. .1618,2146. .2924)
                     /gene="iso-1b"
                     /codon_start=3
                     /product="isoamylase"
                     /protein_id="CAC41016.2"
                     /db_xref="GI:17932913"
                     /translation="SGPAPRLRRWRPNATAGKGVGEVCAAVVEAATKVEDEGEEDEPV
                     AEDRYALGGACRVLAGMPAPLGATALAGGVNFAVYSGGATAAALCLFTPEDLKADRVT
                     EEVPLDPLMNRTGNVWHVFIEGELHNMLYGYRFDGTFAPHCGHYLDVSNVVVDPYAKA
                     VISRGEYGVPARGNNCWPQMAGMIPLPYSTFDWEGDLPLRYPQKDLVIYEMHLRGFTK
                     HDSSNVEHPGTFIGAVSKLDYLKELGVNCIELMPCHEFNELEYSTSSSKMNFWGYSTI
                     NFFSPMTRYTSGGIKNCGRDAINEFKTFVREAHKRGIEVILDVVFNHTAEGNENGPIL
                     SFKGVDNTTYYMLAPKGEFYNYSGCGNTFNCNHPVVRQFIVDCLRYWVTEMHVDGFRF
                     DLASIMTRGSSLWDPVNVYGAPIEGDMITTGTPLVTPPLIDMISNDPILGGVKLIAEA
                     WDAGGLYQVGQFPHWNVWSEWNGKYRDIVRQFIKGTDGFAGGFAECLCGSPHLYQAGG
                     RKPWHSINFVCAHDGFTLADLVTYNKKYNLPNGENNRDGENHNLSWNCGEEGEFARLS
                     VKRLRKRQMRNFFVCLMVSQGVPMFYMGDEYGHTKGGNNNTYCHDSYVNYFRWDKKEQ
                     YSELHRFCCLMTKFRKECEGLGLEDFPTAKRLQWHGHQPGKPDWSENSRFVAFSMKDE
                     RQGEIYVAFNTSHLPAVVELPERAGRRWEPVVDTGKPAPYDFLTDDLPDRALTIHQFS
                     HFLYSNLYPMLSYSSVILVLRPDV"
     intron          298. .397
                     /gene="iso-1b"
                     /number=1
     exon            398. .1618
                     /gene="iso-1b"
                     /number=2
     intron          1619. .2145
                     /gene="iso-1b"
                     /number=2
     exon            2146. .2924
                     /gene="iso-1b"
                     /number=3
BASE COUNT      803 a    636 c    762 g    796 t
ORIGIN


  Query Match           100.0%;  Score 2997;  DB 8;  Length 2997;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 2997;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 ggtcggggccggcgccgcgcctgcgacggtggcgacccaatgcgacggcggggaaggggg 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GGTCGGGGCCGGCGCCGCGCCTGCGACGGTGGCGACCCAATGCGACGGCGGGGAAGGGGG 60

Qy     61 tcggcgaggtgtgcgccgcggttgtcgaggcggcgacgaaggtagaggacgaggggggagg 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TCGGCGAGGTGTGCGCCGCGGTTGTCGAGGCGGCGACGAAGGTAGAGGACGAGGGGGAGG 120

Qy    121 aggacgagccggtggcggaggacaggtacgcgctcggcggcgcgtgcagggtgctcgccg 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 AGGACGAGCCGGTGGCGGAGGACAGGTACGCGCTCGGCGGCGCGTGCAGGGTGCTCGCCG 180

Qy    181 gaatgcccgcgccgctgggcgccaccgcgctcgccggcggggtcaatttcgccgtctatt 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GAATGCCCGCGCCGCTGGGCGCCACCGCGCTCGCCGGCGGGGTCAATTTCGCCGTCTATT 240

Qy    241 ccggcggagccaccgccgcggcgctctgcctcttcacgccagaagatctcaaggcggtgg 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CCGGCGGAGCCACCGCCGCGGCGCTCTGCCTCTTCACGCCAGAAGATCTCAAGGCGGTGG 300

Qy    301 ggttgcctcccgagtagagttcatcagctttgcgtgcgccgcgcgccccttttttgggcc 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GGTTGCCTCCCGAGTAGAGTTCATCAGCTTTGCGTGCGCCGCGCGCCCCTTTTTTGGGCC 360

Qy    361 tgcaatttaagttttgtactggggcaaatgctgcaggatagggtgaccgaggaggttccc 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TGCAATTTAAGTTTTGTACTGGGGCAAATGCTGCAGGATAGGGTGACCGAGGAGGTTCCC 420

Qy    421 cttgacccccctgatgaatcggaccgggaacgtgtggcatgtcttcatcgaaggcgagctg 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 CTTGACCCCCTGATGAATCGGACCGGGAACGTGTGGCATGTCTTCATCGAAGGCGAGCTG 480

Qy    481 cacaacatgctttacgggtacaggttcgacggcacctttgctcctcactgcgggcactac 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CACAACATGCTTTACGGGTACAGGTTCGACGGCACCTTTGCTCCTCACTGCGGGCACTAC 540

Qy    541 cttgatgtttccaatgtcgtggtggatccttatgctaaggcagtgataagccgaggggag 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CTTGATGTTTCCAATGTCGTGGTGGATCCTTATGCTAAGGCAGTGATAAGCCGAGGGGAG 600
```

```
Qy    601 tatggtgttccagcgcgtggtaacaattgctggcctcagatggctggcatgatccctctt 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 TATGGTGTTCCAGCGCGTGGTAACAATTGCTGGCCTCAGATGGCTGGCATGATCCCTCTT 660

Qy    661 ccatatagcacgtttgattgggaaggcgacctacctctaagatatcctcaaaaggacctg 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 CCATATAGCACGTTTGATTGGGAAGGCGACCTACCTCTAAGATATCCTCAAAAGGACCTG 720

Qy    721 gtaatatatgagatgcacttgcgtggattcacgaagcatgattcaagcaatgtagaacat 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GTAATATATGAGATGCACTTGCGTGGATTCACGAAGCATGATTCAAGCAATGTAGAACAT 780

Qy    781 ccgggtactttcattggagctgtgtcgaagcttgactatttgaaggagcttggagttaat 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 CCGGGTACTTTCATTGGAGCTGTGTCGAAGCTTGACTATTTGAAGGAGCTTGGAGTTAAT 840

Qy    841 tgtattgaattaatgccctgccatgagttcaacgagctggagtactcaacctcttcttcc 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 TGTATTGAATTAATGCCCTGCCATGAGTTCAACGAGCTGGAGTACTCAACCTCTTCTTCC 900

Qy    901 aagatgaacttttggggatattctaccataaaacttcttttcaccaatgacaagatacaca 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 AAGATGAACTTTTGGGGATATTCTACCATAAACTTCTTTTCACCAATGACAAGATACACA 960

Qy    961 tcaggcgggataaaaaaactgtgggcgtgatgccataaatgagttcaaaacttttgtaaga 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 TCAGGCGGGATAAAAAACTGTGGGCGTGATGCCATAAATGAGTTCAAAACTTTTGTAAGA 1020

Qy   1021 gaggctcacaaacggggaattgaggtgatcctggatgttgtcttcaaccatacagctgag 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 GAGGCTCACAAACGGGGAATTGAGGTGATCCTGGATGTTGTCTTCAACCATACAGCTGAG 1080

Qy   1081 ggtaatgagaatggtccaatattatcatttaaggggtcgataatactacatactatatg 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 GGTAATGAGAATGGTCCAATATTATCATTTAAGGGGGTCGATAATACTACATACTATATG 1140

Qy   1141 cttgcacccaagggagagttttataactattctggctgtgggaatacctcaactgtaat 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 CTTGCACCCAAGGGAGAGTTTTATAACTATTCTGGCTGTGGGAATACCTTCAACTGTAAT 1200

Qy   1201 catcctgtggttcgtcaattcattgtagattgtttaagatactgggtgacggaaatgcat 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 CATCCTGTGGTTCGTCAATTCATTGTAGATTGTTTAAGATACTGGGTGACGGAAATGCAT 1260

Qy   1261 gttgatggttttcgttttgatcttgcatccataatgaccagaggttccagtctgtgggat 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 GTTGATGGTTTTCGTTTTGATCTTGCATCCATAATGACCAGAGGTTCCAGTCTGTGGGAT 1320

Qy   1321 ccagttaacgtgtatggagctccaatagaaggtgacatgatcacaacagggacacctctt 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 CCAGTTAACGTGTATGGAGCTCCAATAGAAGGTGACATGATCACAACAGGGACACCTCTT 1380

Qy   1381 gttactccaccacttattgacatgatcagcaatgacccaattcttggaggcgtcaagctc 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 GTTACTCCACCACTTATTGACATGATCAGCAATGACCCAATTCTTGGAGGCGTCAAGCTC 1440

Qy   1441 attgctgaagcatgggatgcaggaggcctctatcaagtaggtcaattccctcactggaat 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 ATTGCTGAAGCATGGGATGCAGGAGGCCTCTATCAAGTAGGTCAATTCCCTCACTGGAAT 1500

Qy   1501 gtttggtctgagtggaatgggaagtaccgggacattgtgcgtcaattcattaaaggcact 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 GTTTGGTCTGAGTGGAATGGGAAGTACCGGGACATTGTGCGTCAATTCATTAAAGGCACT 1560

Qy   1561 gatggatttgctggtggttttgccgaatgtctttgtggaagtccacacctataccaggta 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 GATGGATTTGCTGGTGGTTTTGCCGAATGTCTTTGTGGAAGTCCACACCTATACCAGGTA 1620

Qy   1621 agttgtggcaatacttgtaaatgagttgagtgaatgtcacctggatttttttatatatacc 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 AGTTGTGGCAATACTTGTAAATGAGTTGAGTGAATGTCACCTGGATTTTTTATATATACC 1680

Qy   1681 acatgatgatacacatctaaatatataacaatcatagtgtatgcatatgcatttggctaa 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 ACATGATGATACACATCTAAATATATAACAATCATAGTGTATGCATATGCATTTGGCTAA 1740

Qy   1741 gaagtattagtgtatacactagtgctatatataggttttaacacccaacttgccaatgaa 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 GAAGTATTAGTGTATACACTAGTGCTATATATAGGTTTTAACACCCAACTTGCCAATGAA 1800

Qy   1801 ggaacatagggctttctagttatcttatttatttgtccggtgaataatccactgaaaaat 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 GGAACATAGGGCTTTCTAGTTATCTTATTTATTTGTCCGGTGAATAATCCACTGAAAAAT 1860

Qy   1861 tccagccatgtcatttttttaggggggagaagaaactatattgatttgccccccctaaaag 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 TCCAGCCATGTCATTTTTTAGGGGGGGAGAAGAAACTATATTGATTTGCCCCCCTAAAAG 1920

Qy   1921 aagccatctcagaattcataggtaagttgcttttctgtaaagaaaggaaaacgacttcat 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 AAGCCATCTCAGAATTCATAGGTAAGTTGCTTTTCTGTAAAGAAAGGAAAACGACTTCAT 1980

Qy   1981 actttctatcggtgctaacttagctcgatgtatatttgtaagatgaatgccaaatttaat 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 ACTTTCTATCGGTGCTAACTTAGCTCGATGTATATTTGTAAGATGAATGCCAAATTTAAT 2040

Qy   2041 ttgtcggataatttgatctgttattcacaaatttctatttggtttctctagaaatcaaac 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 TTGTCGGATAATTTGATCTGTTATTCACAAATTTCTATTTGGTTTCTCTAGAAATCAAAC 2100

Qy   2101 cagtaacttgttattggcactgcaacttcttattgattaatcaggcaggaggaaggaaac 2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 CAGTAACTTGTTATTGGCACTGCAACTTCTTATTGATTAATCAGGCAGGAGGAAGGAAAC 2160

Qy   2161 cttggcacagtatcaactttgtatgtgcacatgatggatttacactggctgatttggtaa 2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 CTTGGCACAGTATCAACTTTGTATGTGCACATGATGGATTTACACTGGCTGATTTGGTAA 2220

Qy   2221 catataataagaagtacaatttaccaaatggggagaacaacagagatggagaaaatcaca 2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 CATATAATAAGAAGTACAATTTACCAAATGGGGAGAACAACAGAGATGGAGAAAATCACA 2280

Qy   2281 atcttagctggaattgtggggaggaaggagaattcgcaagattgtctgtcaaaagattga 2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 ATCTTAGCTGGAATTGTGGGGAGGAAGGAGAATTCGCAAGATTGTCTGTCAAAAGATTGA 2340

Qy   2341 ggaagaggcagatgcgcaatttctttgtttgtctcatggtttctcaaggagttccaatgt 2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 GGAAGAGGCAGATGCGCAATTTCTTTGTTTGTCTCATGGTTTCTCAAGGAGTTCCAATGT 2400

Qy   2401 tctacatgggtgatgaatatggccacacaaaaggggcaacaacaatacatactgccatg 2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401 TCTACATGGGTGATGAATATGGCCACACAAAAGGGGCAACAACAATACATACTGCCATG 2460

Qy   2461 attcttatgtcaattatttttcgctgggataaaaaagaacaatactctgagttgcaccgat 2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2461 ATTCTTATGTCAATTATTTTCGCTGGGATAAAAAAGAACAATACTCTGAGTTGCACCGAT 2520

Qy   2521 tctgctgcctcatgaccaaattccgcaaggagtgcgagggtcttggccttgaggactttc 2580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2521 TCTGCTGCCTCATGACCAAATTCCGCAAGGAGTGCGAGGGTCTTGGCCTTGAGGACTTTC 2580

Qy   2581 caacggccaaacggctgcagtggcatggtcatcagcctgggaagcctgattggtctgaga 2640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2581 CAACGGCCAAACGGCTGCAGTGGCATGGTCATCAGCCTGGGAAGCCTGATTGGTCTGAGA 2640

Qy   2641 atagccgattcgttgccttttccatgaaagatgaaagacagggcgagatctatgtggcct 2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2641 ATAGCCGATTCGTTGCCTTTTCCATGAAAGATGAAAGACAGGGCGAGATCTATGTGGCCT 2700

Qy   2701 tcaacaccagccacttaccggccgttgttgagctcccagagcgcgcagggcgccggtggg 2760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2701 TCAACACCAGCCACTTACCGGCCGTTGTTGAGCTCCCAGAGCGCGCAGGGCGCCGGTGGG 2760

Qy   2761 aaccggtggtggacacaggcaagccagcaccatacgacttcctcaccgacgacttacctg 2820
```

```
Db    2761 AACCGGTGGTGGACACAGGCAAGCCAGCACCATACGACTTCCTCACCGACGACTTACCTG 2820

Qy    2821 atcgcgctctcaccatacaccagttctcgcatttcctctactccaacctctaccccatgc 2880

Db    2821 ATCGCGCTCTCACCATACACCAGTTCTCGCATTTCCTCTACTCCAACCTCTACCCCATGC 2880

Qy    2881 tcagctactcatcggtcatcctagtattgcgccctgatgtttgagagaccaatatatataca 2940

Db    2881 TCAGCTACTCATCGGTCATCCTAGTATTGCGCCCTGATGTTTGAGAGACCAATATATATACA 2940

Qy    2941 gtaaataatatgtctatatgtaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 2997

Db    2941 GTAAATAATATGTCTATATGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2997


RESULT   4
AF438328
LOCUS       AF438328               2590 bp    mRNA    linear   PLN 13-NOV-2001
DEFINITION  Triticum aestivum isoamylase precursor, mRNA, complete cds;
            alternatively spliced.
ACCESSION   AF438328
VERSION     AF438328.1  GI:16905062
KEYWORDS    .
SOURCE      bread wheat.
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Triticum.
REFERENCE   1  (bases 1 to 2590)
  AUTHORS   Netrphan,S., Khandelwal,R.L. and Chibbar,R.N.
  TITLE     Characterization of isoamylase transcripts produced during wheat
            grain development and germination
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2590)
  AUTHORS   Netrphan,S., Khandelwal,R.L. and Chibbar,R.N.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-OCT-2001) Plant Biotechnology Institute, 110
            Gymnasium Place, Saskatoon, SK S7N 0W9, Canada
FEATURES             Location/Qualifiers
     source          1. .2590
                     /organism="Triticum aestivum"
                     /cultivar="Fielder"
                     /db_xref="taxon:4565"
                     /tissue_type="endosperm"
     CDS             14. .2386
                     /EC_number="3.2.1.68"
                     /function="hydrolyzes alpha-1,6-glucosidic bonds of
                     starch, amylopectin, and glycogen"
                     /note="alternatively spliced; glycogen 6-glucanohydrolase"
                     /codon_start=1
                     /product="isoamylase precursor"
                     /protein_id="AAL31015.1"
                     /db_xref="GI:16905063"
                     /translation="MTMMAMPEAPCLCARPSLAARARRPGSGPAPRLRRWRPNATAGK
                     GVGEVCAAVVEAATKVEDEGEEDEPVAEDRYALGGACRVLAGMPAPLGATALAGGVNF
                     AVYSGGATAAALCLFTPEDLKADRVTEEVPLDPLMNRTGNVWHVFIEGELHNMLYGYR
                     FDGTFAPHCGHYLDVSNVVVDPYAKAVISRGEYGVPARGNNCWPQMAGMIPLPYSTFD
                     WEGDLPLRYPQKDLVIYEMHLRGFTKHDSSNVEHPGTFIGAVSKLDYLKELGVNCIEL
                     MPCHEFNELEYSTSSSKMNFWGYSTINFFSPMTRYTSGGIKNCGRDAINEFKTFVREA
                     HKRGIEVILDVVFNHTAEGNENGPILSFRGVDNTTYYMLAPKGEFYNYSGCGNTFNCN
                     HPVVRQFIVDCLRYWVTEMHVDGFRFDLASIMTRGSSLWDPVNVYGAPIEGDMITTGT
                     PLVTPPLIDMISNDPILGGVKLIAEAWDAGGLYQVGQFPHWNVWSEWNGKYRDIVRQF
                     IKGTDGFAGGFAECLCGSPHLYQAGGRKPWHSINFVCAHDGFTLADLVTYNKKYNLPN
                     GENNRDGENHNLSWNCGEEGEFARLSVKRLRKRQMRNFFVCLMVSQGVPMFYMGDEYG
                     HTKGGNNNIYCHDSYVNYFRWDKKEQYSELHRFCCLMTKFRKECEGLGLEDFPTAKRL
                     QWHGHQPGKPDWSENSRFVAFSMKDERQGEIYVAFNTSHLPAVVELPERAGRRWEPVV
                     DTGKPAPYDFLTDDLPDRALTIHQFSHFLYSNLYPMLSYSSVILVLRPDV"
     mat_peptide     167. .2383
                     /product="isoamylase"
     polyA_signal    2406. .2411
BASE COUNT      676 a    587 c    695 g    632 t
ORIGIN
```

```
Query Match           56.2%;  Score 1684.8;  DB 8;  Length 2590;
Best Local Similarity  78.8%;  Pred. No. 4.3e-306;
Matches 2333;  Conservative    0;  Mismatches    2;  Indels  627;  Gaps    2;

Qy      1 ggtcggggccggcgccgcgcctgcgacggtggcgacccaatgcgacggcggggaaggggg 60

Db     90 GGTCGGGGCCGGCGCCGCGCCTGCGACGGTGGCGACCCAATGCGACGGCGGGGAAGGGGG 149

Qy     61 tcggcgaggtgtgcgccgcggttgtcgaggcggcgacgaaggtagaggacgaggggggagg 120

Db    150 TCGGCGAGGTGTGCGCCGCGGTTGTCGAGGCGGCGACGAAGGTAGAGGACGAGGGGGAGG 209

Qy    121 aggacgagccggtggcggaggacaggtacgcgctcggcggcgcgtgcagggtgctcgccg 180

Db    210 AGGACGAGCCGGTGGCGGAGGACAGGTACGCGCTCGGCGGCGCGTGCAGGGTGCTCGCCG 269

Qy    181 gaatgcccgcgccgctgggcgccaccgcgctcgccggcggggtcaatttcgccgtctatt 240

Db    270 GAATGCCCGCGCCGCTGGGCGCCACCGCGCTCGCCGGCGGGGTCAATTTCGCCGTCTATT 329

Qy    241 ccggcggagccaccgccgcggcgctctgcctcttcacgccagaagatctcaaggcggtgg 300

Db    330 CCGGCGGAGCCACCGCCGCGGCGCTCTGCCTCTTCACGCCAGAAGATCTCAAGGC----- 384

Qy    301 ggttgcctcccgagtagagttcatcagctttgcgtgcgccgcgcgccccttttttgggcc 360

Db    385 ------------------------------------------------------------ 384

Qy    361 tgcaatttaagttttgtactggggcaaatgctgcaggatagggtgaccgaggaggttccc 420

Db    385 ----------------------------------GGATAGGGTGACCGAGGAGGTTCCC 409

Qy    421 cttgacccccctgatgaatcggaccgggaacgtgtggcatgtcttcatcgaaggcgagctg 480

Db    410 CTTGACCCCCTGATGAATCGGACCGGGAACGTGTGGCATGTCTTCATCGAAGGCGAGCTG 469

Qy    481 cacaacatgctttacgggtacaggttcgacggcacctttgctcctcactgcgggcactac 540

Db    470 CACAACATGCTTTACGGGTACAGGTTCGACGGCACCTTTGCTCCTCACTGCGGGCACTAC 529

Qy    541 cttgatgtttccaatgtcgtggtggatccttatgctaaggcagtgataagccgaggggag 600

Db    530 CTTGATGTTTCCAATGTCGTGGTGGATCCTTATGCTAAGGCAGTGATAAGCCGAGGGGAG 589

Qy    601 tatggtgttccagcgcgtggtaacaattgctggcctcagatggctggcatgatccctctt 660

Db    590 TATGGTGTTCCAGCGCGTGGTAACAATTGCTGGCCTCAGATGGCTGGCATGATCCCTCTT 649

Qy    661 ccatatagcacgtttgattgggaaggcgacctacctctaagatatcctcaaaaggacctg 720

Db    650 CCATATAGCACGTTTGATTGGGAAGGCGACCTACCTCTAAGATATCCTCAAAAGGACCTG 709

Qy    721 gtaatatatgagatgcacttgcgtggattcacgaagcatgattcaagcaatgtagaacat 780

Db    710 GTAATATATGAGATGCACTTGCGTGGATTCACGAAGCATGATTCAAGCAATGTAGAACAT 769

Qy    781 ccgggtactttcattggagctgtgtcgaagcttgactatttgaaggagcttggagttaat 840

Db    770 CCGGGTACTTTCATTGGAGCTGTGTCGAAGCTTGACTATTTGAAGGAGCTTGGAGTTAAT 829

Qy    841 tgtattgaattaatgccctgccatgagttcaacgagctggagtactcaacctcttcttcc 900

Db    830 TGTATTGAATTAATGCCCTGCCATGAGTTCAACGAGCTGGAGTACTCAACCTCTTCTTCC 889

Qy    901 aagatgaacttttggggatattctaccataaacttcttttcaccaatgacaagatacaca 960

Db    890 AAGATGAACTTTTGGGGATATTCTACCATAAACTTCTTTTCACCAATGACAAGATACACA 949

Qy    961 tcaggcgggataaaaaactgtgggcgtgatgccataaatgagttcaaaacttttgtaaga 1020

Db    950 TCAGGCGGGATAAAAAACTGTGGGCGTGATGCCATAAATGAGTTCAAAACTTTTGTAAGA 1009
```

```
Qy   1021 gaggctcacaaacggggaattgaggtgatcctggatgttgtcttcaaccatacagctgag 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1010 GAGGCTCACAAACGGGGAATTGAGGTGATCCTGGATGTTGTCTTCAACCATACAGCTGAG 1069

Qy   1081 ggtaatgagaatggtccaatattatcatttaaggggtcgataatactacatactatatg 1140
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db   1070 GGTAATGAGAATGGTCCAATATTATCATTTAGGGGGGTCGATAATACTACATACTATATG 1129

Qy   1141 cttgcacccaagggagagttttataactattctggctgtgggaatacctttcaactgtaat 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1130 CTTGCACCCAAGGGAGAGTTTTATAACTATTCTGGCTGTGGGAATACCTTCAACTGTAAT 1189

Qy   1201 catcctgtggttcgtcaattcattgtagattgtttaagatactgggtgacggaaatgcat 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1190 CATCCTGTGGTTCGTCAATTCATTGTAGATTGTTTAAGATACTGGGTGACGGAAATGCAT 1249

Qy   1261 gttgatggttttcgttttgatcttgcatccataatgaccagaggttccagtctgtgggat 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1250 GTTGATGGTTTTCGTTTTGATCTTGCATCCATAATGACCAGAGGTTCCAGTCTGTGGGAT 1309

Qy   1321 ccagttaacgtgtatggagctccaatagaaggtgacatgatcacaacagggacacctctt 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1310 CCAGTTAACGTGTATGGAGCTCCAATAGAAGGTGACATGATCACAACAGGGACACCTCTT 1369

Qy   1381 gttactccaccacttattgacatgatcagcaatgacccaattcttggaggcgtcaagctc 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1370 GTTACTCCACCACTTATTGACATGATCAGCAATGACCCAATTCTTGGAGGCGTCAAGCTC 1429

Qy   1441 attgctgaagcatgggatgcaggaggcctctatcaagtaggtcaattccctcactggaat 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1430 ATTGCTGAAGCATGGGATGCAGGAGGCCTCTATCAAGTAGGTCAATTCCCTCACTGGAAT 1489

Qy   1501 gtttggtctgagtggaatgggaagtaccgggacattgtgcgtcaattcattaaaggcact 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1490 GTTTGGTCTGAGTGGAATGGGAAGTACCGGGACATTGTGCGTCAATTCATTAAAGGCACT 1549

Qy   1561 gatggatttgctggtggtttgccgaatgtctttgtggaagtccacacctataccaggta 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1550 GATGGATTTGCTGGTGGTTTTGCCGAATGTCTTTGTGGAAGTCCACACCTATAC------ 1603

Qy   1621 agttgtggcaatacttgtaaatgagttgagtgaatgtcacctggatttttatatatacc 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1604 ---------------------------------------------------------- 1603

Qy   1681 acatgatgatacacatctaaatatataacaatcatagtgtatgcatatgcatttggctaa 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1604 ---------------------------------------------------------- 1603

Qy   1741 gaagtattagtgtatacactagtgctatatataggttttaacacccaacttgccaatgaa 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1604 ---------------------------------------------------------- 1603

Qy   1801 ggaacatagggctttctagttatcttatttatttgtccggtgaataatccactgaaaaat 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1604 ---------------------------------------------------------- 1603

Qy   1861 tccagccatgtcatttttaggggggagaagaaactatattgatttgccccctaaaag 1920
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1604 ---------------------------------------------------------- 1603

Qy   1921 aagccatctcagaattcataggtaagttgctttttctgtaaagaaaggaaaacgacttcat 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1604 ---------------------------------------------------------- 1603

Qy   1981 actttctatcggtgctaacttagctcgatgtatatttgtaagatgaatgccaaatttaat 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1604 ---------------------------------------------------------- 1603

Qy   2041 ttgtcggataatttgatctgttattcacaaatttctatttggtttctctagaaatcaaac 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1604 ---------------------------------------------------------- 1603

Qy   2101 cagtaacttgttattggcactgcaacttcttattgattaatcaggcaggaggaaggaaac 2160
          |||||||||||||||||||                                        ||||||||||||||||||
Db   1604 ----------------------------------------CAGGCAGGAGGAAGGAAAC 1622

Qy   2161 cttggcacagtatcaactttgtatgtgcacatgatggatttacactggctgatttggtaa 2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1623 CTTGGCACAGTATCAACTTTGTATGTGCACATGATGGATTTACACTGGCTGATTTGGTAA 1682

Qy   2221 catataataagaagtacaatttaccaaatggggagaacaacagagatggagaaaatcaca 2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1683 CATATAATAAGAAGTACAATTTACCAAATGGGGAGAACAACAGAGATGGAGAAAATCACA 1742

Qy   2281 atcttagctggaattgtggggaggaaggagaattcgcaagattgtctgtcaaaagattga 2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1743 ATCTTAGCTGGAATTGTGGGGAGGAAGGAGAATTCGCAAGATTGTCTGTCAAAAGATTGA 1802

Qy   2341 ggaagaggcagatgcgcaatttctttgtttgtctcatggtttctcaaggagttccaatgt 2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1803 GGAAGAGGCAGATGCGCAATTTCTTTGTTTGTCTCATGGTTTCTCAAGGAGTTCCAATGT 1862

Qy   2401 tctacatgggtgatgaatatggccacacaaaaggggcaacaacaatacatactgccatg 2460
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1863 TCTACATGGGTGATGAATATGGCCACACAAAAGGGGCAACAACAATATATACTGCCATG 1922

Qy   2461 attcttatgtcaattattttcgctgggataaaaaaagaacaatactctgagttgcaccgat 2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1923 ATTCTTATGTCAATTATTTTCGCTGGGATAAAAAAGAACAATACTCTGAGTTGCACCGAT 1982

Qy   2521 tctgctgcctcatgaccaaattccgcaaggagtgcgagggtcttggccttgaggactttc 2580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1983 TCTGCTGCCTCATGACCAAATTCCGCAAGGAGTGCGAGGGTCTTGGCCTTGAGGACTTTC 2042

Qy   2581 caacggccaaacggctgcagtggcatggtcatcagcctgggaagcctgattggtctgaga 2640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2043 CAACGGCCAAACGGCTGCAGTGGCATGGTCATCAGCCTGGGAAGCCTGATTGGTCTGAGA 2102

Qy   2641 atagccgattcgttgccttttccatgaaagatgaaagacagggcgagatctatgtggcct 2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2103 ATAGCCGATTCGTTGCCTTTTCCATGAAAGATGAAAGACAGGGCGAGATCTATGTGGCCT 2162

Qy   2701 tcaacaccagccacttaccggccgttgttgagctcccagagcgcgcagggcgccggtggg 2760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2163 TCAACACCAGCCACTTACCGGCCGTTGTTGAGCTCCCAGAGCGCGCAGGGCGCCGGTGGG 2222

Qy   2761 aaccggtggtggacacaggcaagccagcaccatacgacttcctcaccgacgacttacctg 2820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2223 AACCGGTGGTGGACACAGGCAAGCCAGCACCATACGACTTCCTCACCGACGACTTACCTG 2282

Qy   2821 atcgcgctctcaccatacaccagttctcgcatttcctctactccaacctctaccccatgc 2880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2283 ATCGCGCTCTCACCATACACCAGTTCTCGCATTTCCTCTACTCCAACCTCTACCCCATGC 2342

Qy   2881 tcagctactcatcggtcatcctagtattgcgccctgatgtttgagagaccaatatatataca 2940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2343 TCAGCTACTCATCGGTCATCCTAGTATTGCGCCCTGATGTTTGAGAGACCAATATATATACA 2402

Qy   2941 gtaaataatatgtctatatgta 2962
          |||||||||||||||||||||
Db   2403 GTAAATAATATGTCTATATGTA 2424
```

```
RESULT   5
TAE301647
LOCUS       TAE301647              2490 bp    mRNA    linear   PLN 17-DEC-2001
DEFINITION  Triticum aestivum mRNA for isoamylase (isol gene).
ACCESSION   AJ301647
VERSION     AJ301647.1  GI:17932897
KEYWORDS    isol gene; isoamylase.
SOURCE      bread wheat.
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Triticum.
```

```
REFERENCE   1  (bases 1 to 2490)
  AUTHORS   Genschel,U.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-NOV-2000) Genschel U., Institut fuer Allgemeine
            Botanik, AMP II, Universitaet Hamburg, Ohnhorststrasse 18, Hamburg,
            22609, Germany
  REMARK    revised by submitter 22-MAR-2001
COMMENT     Related entry AJ307689.
FEATURES             Location/Qualifiers
     source          1. .2490
                     /organism="Triticum aestivum"
                     /cultivar="Florida"
                     /db_xref="taxon:4565"
                     /tissue_type="developing kernel"
     transit_peptide 6. .214
                     /gene="isol"
     CDS             6. .2363
                     /gene="isol"
                     /standard_name="sugary-1 isoamylase"
                     /codon_start=1
                     /evidence=experimental
                     /product="isoamylase"
                     /protein_id="CAC82925.1"
                     /db_xref="GI:17932898"
                     /translation="MAKAPCLCARPSLAARTRRPGPGPAPRLRRWRPNATAGKGVGEV
                     CAAVVEVATKAEDEGEEDEPVAEDRYALGGACRVLAGMPTPLGATALAGGVNFAVYSG
                     GATAAALCLFTPEDLKADRVTEEVPLDPLMNRTGNVWHVFIEGELQDMLYGYRFDGTF
                     APHCGHYLDVSNVVVDPYAKAVISRGEYGVPARGNNCWPQMAGMIPLPYSTFDWEGDL
                     PLRYPQKDLVIYEMHLRGFTKHDSSNVEHPGTFIGAVSKLDYLKELGVNCIELMPCHE
                     FNELEYSTSSSKMNFWGYSTINFFSPMTRYTSGGIKNCGRDAINEFKTFVREAHKRGI
                     EVILDVVFNHTAEGNENGPILSFRGVDNTTYYMLAPKGEFYNYSGCGNTFNCNHPVVR
                     QFIVDCLRYWVTEMHVDGFRFDLASIMTRGSSLWDPVNVYGAPIEGDMITTGTPLVTP
                     PLIDMISNDPILGGVKLVAEAWDAGGLYQVGQFPHWNVWSEWNGKYRDIVRQFIKGTD
                     GFAGGFAECLCGSPHLYQAGGRKPWHSINFVCAHDGFTLADLVTYNNKYNLPNGENNR
                     DGENHNLSWNCGEEGEFARLSVKRLRKRQMRNFFVCLMVSQGVPMFYMGDEYGHTKGG
                     NNNTYCHDSYVNYFRWDKKEQYSDLHRFCCLMTKFRKECEGLGLEDFPTAERLQWHGH
                     QPGKPDWSENSRFVAFSMKDERQGEIYVAFNTSHLPAVVELPERTGRRWEPVVDTGKP
                     APYDFLTDDLPDRALTIHQFSHFLNSNLYPMLSYSSVILVLRPDV"
     gene            6. .2363
                     /gene="isol"
     mat_peptide     215. .2360
                     /gene="isol"
                     /product="isoamylase"
BASE COUNT      619 a    573 c    683 g    615 t
ORIGIN
```

Query Match 53.9%; Score 1614.8; DB 8; Length 2490;
Best Local Similarity 77.5%; Pred. No. 5.9e-293;
Matches 2296; Conservative 0; Mismatches 37; Indels 629; Gaps 3;

```
Qy       1 ggtcggggccggcgccgcgcctgcgacggtggcgacccaatgcgacggcggggaaggggg 60
Db      67 GGCCGGGGCCGGCGCCGCGCCTGCGACGGTGGCGGCCCAATGCGACGGCGGGGAAGGGGG 126

Qy      61 tcggcgaggtgtgcgccgcggttgtcgaggcggcgacgaaggtagaggacgaggggggagg 120
Db     127 TCGGCGAGGTCTGCGCCGCGGTTGTCGAGGTGGCGACGAAGGCCGAGGATGAGGGGGAGG 186

Qy     121 aggacgagccggtggcggaggacaggtacgcgctcggcggcgcgtgcagggtgctcgccg 180
Db     187 AGGACGAGCCGGTGGCGGAGGACAGGTACGCGCTCGGCGGCGCGTGCAGGGTGCTCGCCG 246

Qy     181 gaatgcccgcgccgctgggcgccaccgcgctcgccggcggggtcaatttcgccgtctatt 240
Db     247 GAATGCCCACGCCGCTGGGCGCCACCGCGCTCGCCGGCGGGGTCAATTTCGCCGTCTACT 306

Qy     241 ccggcggagccaccgccgcggcgctctgcctcttcacgccagaagatctcaaggcggtgg 300
Db     307 CCGGCGGAGCCACAGCCGCGGCGCTCTGCCTCTTCACGCCAGAAGATCTCAAGGC----- 361

Qy     301 ggttgcctcccgagtagagttcatcagctttgcgtgcgccgcgcgccccttttttgggcc 360
```

```
Db     362 ---------------------------------------------------------- 361

Qy     361 tgcaatttaagttttgtactggggcaaatgctgcaggatagggtgaccgaggaggttccc 420
Db     362 ---------------------------------GGATAGGGTGACGGAGGAGGTTCCC 386

Qy     421 cttgacccccctgatgaatcggaccgggaacgtgtggcatgtcttcatcgaaggcgagctg 480
Db     387 CTTGACCCCCTGATGAATCGGACTGGGAACGTATGGCATGTCTTCATCGAAGGCGAGCTG 446

Qy     481 cacaacatgctttacgggtacaggttcgacggcacctttgctcctcactgcgggcactac 540
Db     447 CAGGACATGCTTTACGGGTACAGGTTCGACGGCACCTTTGCTCCTCACTGCGGGCACTAC 506

Qy     541 cttgatgtttccaatgtcgtggtggatccttatgctaaggcagtgataagccgaggggag 600
Db     507 CTTGATGTTTCCAATGTCGTGGTGGATCCTTATGCTAAGGCAGTGATAAGCCGAGGGGAG 566

Qy     601 tatggtgttccagcgcgtggtaacaattgctggcctcagatggctggcatgatccctctt 660
Db     567 TATGGTGTTCCGGCGCGTGGTAACAATTGCTGGCCTCAGATGGCTGGCATGATCCCTCTT 626

Qy     661 ccatatagcacgtttgattgggaaggcgacctacctctaagatatcctcaaaaggacctg 720
Db     627 CCATATAGCACGTTTGATTGGGAAGGCGACCTACCTCTAAGATATCCTCAAAAGGACCTG 686

Qy     721 gtaatatatgagatgcacttgcgtggattcacgaagcatgattcaagcaatgtagaacat 780
Db     687 GTAATATATGAGATGCACTTGCGTGGATTCACGAAGCATGATTCAAGCAATGTAGAACAT 746

Qy     781 ccgggtactttcattggagctgtgtcgaagcttgactatttgaaggagcttggagttaat 840
Db     747 CCCGGTACTTTCATTGGGGCTGTGTCGAAGCTTGACTATTTGAAGGAGCTTGGAGTTAAT 806

Qy     841 tgtattgaattaatgccctgccatgagttcaacgagctggagtactcaacctcttcttcc 900
Db     807 TGTATTGAGTTAATGCCCTGCCATGAGTTCAACGAGCTGGAGTACTCAACCTCTTCTTCC 866

Qy     901 aagatgaacttttggggatattctaccataaacttcttttcaccaatgacaagatacaca 960
Db     867 AAGATGAACTTTTGGGGATATTCTACCATAAACTTCTTTTCACCAATGACGAGATACACA 926

Qy     961 tcaggcgggataaaaaaactgtgggcgtgatgccataaatgagttcaaaacttttgtaaga 1020
Db     927 TCAGGCGGGATAAAAAACTGTGGGCGTGATGCCATAAATGAGTTCAAAACTTTTGTAAGA 986

Qy    1021 gaggctcacaaacggggaattgaggtgatcctggatgttgtcttcaaccatacagctgag 1080
Db     987 GAGGCTCACAAACGGGGAATTGAGGTGATCCTGGATGTTGTCTTCAACCATACAGCTGAG 1046

Qy    1081 ggtaatgagaatggtccaatattatcatttaaggggggtcgataatactacatactatatg 1140
Db    1047 GGTAATGAGAATGGTCCAATATTATCATTTAGGGGGGTCGATAATACTACATACTATATG 1106

Qy    1141 cttgcacccaagggagagttttataactattctggctgtgggaatacctttcaactgtaat 1200
Db    1107 CTTGCACCCAAGGGAGAGTTTTATAACTATTCTGGCTGTGGGAATACCTTCAACTGTAAT 1166

Qy    1201 catcctgtggttcgtcaattcattgtagattgtttaagatactgggtgacggaaatgcat 1260
Db    1167 CATCCTGTGGTTCGTCAATTCATTGTAGATTGTTTAAGATACTGGGTGACGGAAATGCAT 1226

Qy    1261 gttgatggttttcgttttgatcttgcatccataatgaccagaggttccagtctgtgggat 1320
Db    1227 GTTGATGGTTTTCGTTTTGATCTTGCATCCATAATGACCAGAGGTTCCAGTCTGTGGGAT 1286

Qy    1321 ccagttaacgtgtatggagctccaatagaaggtgacatgatcacaacagggacacctctt 1380
Db    1287 CCAGTTAACGTGTATGGAGCTCCAATAGAAGGTGACATGATCACAACAGGGACACCTCTT 1346

Qy    1381 gttactccaccacttattgacatgatcagcaatgacccaattcttggaggcgtcaagctc 1440
Db    1347 GTTACTCCACCACTTATTGACATGATCAGCAATGACCCAATTCTTGGAGGCGTCAAGCTC 1406
```

```
Qy   1441 attgctgaagcatgggatgcaggaggcctctatcaagtaggtcaattccctcactggaat 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1407 GTTGCTGAAGCATGGGATGCAGGAGGCCTCTATCAAGTAGGTCAATTCCCTCACTGGAAT 1466

Qy   1501 gtttggtctgagtggaatgggaagtaccgggacattgtgcgtcaattcattaaaggcact 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1467 GTTTGGTCTGAGTGGAATGGGAAGTACCGGGACATTGTGCGTCAATTCATTAAAGGCACT 1526

Qy   1561 gatggatttgctggtggttttgccgaatgtctttgtggaagtccacacctataccaggta 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1527 GATGGATTTGCTGGTGGTTTTGCCGAATGTCTTTGTGGAAGTCCACACCTATAC------ 1580

Qy   1621 agttgtggcaatacttgtaaatgagttgagtgaatgtcacctggatttttatatatacc 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1581 ---------------------------------------------------------- 1580

Qy   1681 acatgatgatacacatctaaatatataacaatcatagtgtatgcatatgcatttggctaa 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1581 ---------------------------------------------------------- 1580

Qy   1741 gaagtattagtgtatacactagtgctatatataggttttaacacccaacttgccaatgaa 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1581 ---------------------------------------------------------- 1580

Qy   1801 ggaacatagggctttctagttatcttatttatttgtccggtgaataatccactgaaaaat 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1581 ---------------------------------------------------------- 1580

Qy   1861 tccagccatgtcatttttttaggggggagaagaaactatattgatttgccccctaaaag 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1581 ---------------------------------------------------------- 1580

Qy   1921 aagccatctcagaattcataggtaagttgcttttctgtaaagaaaggaaaacgacttcat 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1581 ---------------------------------------------------------- 1580

Qy   1981 actttctatcggtgctaacttagctcgatgtatatttgtaagatgaatgccaaatttaat 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1581 ---------------------------------------------------------- 1580

Qy   2041 ttgtcggataatttgatctgttattcacaaatttctatttggtttctctagaaatcaaac 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1581 ---------------------------------------------------------- 1580

Qy   2101 cagtaacttgttattggcactgcaacttcttattgattaatcaggcaggaggaaggaaac 2160
                                                 ||||||||||||||||||||
Db   1581 ------------------------------------------CAGGCAGGAGGAAGGAAAC 1599

Qy   2161 cttggcacagtatcaactttgtatgtgcacatgatggatttacactggctgatttggtaa 2220
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db   1600 CTTGGCACAGTATCAACTTTGTATGTGCACACGATGGATTTACACTGGCTGATTTGGTAA 1659

Qy   2221 catataataagaagtacaatttaccaaatggggagaacaacagagatggagaaaatcaca 2280
          |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1660 CATATAATAACAAGTACAATTTACCAAATGGGGAGAACAACAGAGATGGAGAAAATCACA 1719

Qy   2281 atcttagctggaattgtggggaggaaggagaattcgcaagattgtctgtcaaaagattga 2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1720 ATCTTAGCTGGAATTGTGGGGAGGAAGGAGAATTCGCAAGATTGTCTGTCAAAAGATTGA 1779

Qy   2341 ggaagaggcagatgcgcaatttctttgtttgtctcatggtttctcaaggagttccaatgt 2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1780 GGAAGAGGCAGATGCGCAATTTCTTTGTTTGTCTCATGGTTTCTCAAGGAGTTCCAATGT 1839

Qy   2401 tctacatgggtgatgaatatggccacacaaaaggggcaacaacaatacatactgccatg 2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1840 TCTACATGGGTGATGAATATGGCCACACAAAAGGGGCAACAACAATACATACTGCCATG 1899

Qy   2461 attcttatgtcaattattttcgctgggataaaaaaagaacaatactctgagttgcaccgat 2520
          |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db   1900 ATTCTTATGTCAATTATTTTCGCTGGGATAAAAAAGAACAATACTCTGACTTGCACCGAT 1959

Qy   2521 tctgctgcctcatgaccaaattccgcaaggagtgcgagggtcttggccttgaggactttc 2580
          |||| ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db   1960 TCTGTTGCCTCATGACCAAATTCCGCAAGGAGTGCGAGGGTCTTGGCCTTGAGGATTTTC 2019

Qy   2581 caacggccaaacggctgcagtggcatggtcatcagcctgggaagcctgattggtctgaga 2640
          ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2020 CAACGGCCGAACGGCTGCAGTGGCATGGTCATCAGCCTGGGAAGCCTGATTGGTCTGAGA 2079

Qy   2641 atagccgattcgttgccttttccatgaaagatgaaagacagggcgagatctatgtggcct 2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2080 ATAGCCGATTCGTTGCCTTTTCCATGAAAGATGAAAGACAGGGCGAGATCTATGTGGCCT 2139

Qy   2701 tcaacaccagccacttaccggccgttgttgagctcccagagcgcgcagggcgccggtggg 2760
          |||||||||||||||||||||||||||||||||||| |||||| ||||||||||||||||
Db   2140 TCAACACCAGCCACTTACCGGCCGTTGTTGAGCTCCCGGAGCGCACAGGGCGCCGGTGGG 2199

Qy   2761 aaccggtggtggacacaggcaagccagcaccatacgacttcctcaccgacgacttacctg 2820
          |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db   2200 AACCGGTGGTGGACACAGGCAAGCCAGCACCATACGACTTCCTCACTGACGACTTACCTG 2259

Qy   2821 atcgcgctctcaccatacaccagttctcgcatttcctctactccaacctctaccccatgc 2880
          |||||||||||||||||||||||||||||||| ||||||||| ||||||||||||||||||||
Db   2260 ATCGCGCTCTCACCATACACCAGTTCTCTCATTTCCTCAACTCCAACCTCTACCCCATGC 2319

Qy   2881 tcagctactcatcggtcatcctagtattgcgccctgatgtttgagagaccaatatatataca 2940
          |||||||||||||||||||||||||||||||||||||||||||||||| |    |||||||
Db   2320 TCAGCTACTCATCGGTCATCCTAGTATTGCGCCCTGATGTTTGAGAGGC--GGATATACA 2377

Qy   2941 gtaaataatatgtctatatgta 2962
          |||||||||||||| ||||||||
Db   2378 GTAAATAATATGTATATATGTA 2399
```

```
RESULT   6
AX010486
LOCUS       AX010486                2437 bp    DNA     linear   PAT 06-SEP-2000
DEFINITION  Sequence 6 from Patent WO9958690.
ACCESSION   AX010486
VERSION     AX010486.1  GI:9997329
KEYWORDS    .
SOURCE      bread wheat.
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Triticum.
REFERENCE   1  (bases 1 to 2437)
  AUTHORS   Luetticke,S., Abel,G., Loerz,H. and Genschel,U.
  TITLE     Nucleic acid molecules which code for enzymes derived from wheat
            and which are involved in the synthesis of starch
  JOURNAL   Patent: WO 9958690-A 6 18-NOV-1999;
            LUETTICKE STEPHANIE (DE); ABEL GERNOT (DE); LOERZ HORST (DE);
            GENSCHEL ULRICH (DE); HOECHST SCHERING AGREVO GMBH (DE)
FEATURES             Location/Qualifiers
     source          1. .2437
                     /organism="Triticum aestivum"
                     /cultivar="Florida"
                     /db_xref="taxon:4565"
     CDS             <16. .2304
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC07515.1"
                     /db_xref="GI:9997330"
                     /translation="PAPRLRRWRPNATAGKGVGEVCAAVVEVATKAEDEGEEDEPVAE
                     DRYALGGACRVLAGMPTPLGATALAGGVNFAVYSGGATAAALCLFTPEDLKADRVTEE
                     VPLDPLMNRTGNVWHVFIEGELQDMLYGYRFDGTFAPHCGHYLDVSNVVVDPYAKAVI
                     SRGEYGVPARGNNCWPQMAGMIPLPYSTFDWEGDLPLRYPQKDLVIYEMHLRGFTKHD
                     SSNVEHPGTFIGAVSKLDYLKELGVNCIELMPCHEFNELEYSTSSSKMNFWGYSTINF
                     FSPMTRYTSGGIKNCGRDAINEFKTFVREAHKRGIEVILDVVFNHTAEGNENGPILSF
                     RGVDNTTYYMLAPKGEFYNYSGCGNTFNCNHPVVRQFIVDCLRYWVTEMHVDGFRFDL
                     ASIMTRGSSLWDPVNVYGAPIEGDMITTGTPLVTPPLIDMISNDPILGGVKLVAEAWD
                     AGGLYQVGQFPHWNVWSEWNGKYRDIVRQFIKGTDGFAGGFAECLCGSPHLYQAGGRK
                     PWHSINFVCAHDGFTLADLVTYNNKYNLPNGENNRDGENHNLSWNCGEEGEFARLSVK
                     RLRKRQMRNFFVCLMVSQGVPMFYMGDEYGHTKGGNNNTYCHDSYVNYFRWDKKEQYS
```

```
                    DLHRFCCLMTKFRKECEGLGLEDFPTAERLQWHGHQPGKPDWSENSRFVAFSMKDERQ
                    GEIYVAFNTSHLPAVVELPERTGRRWEPVVDTGKPAPYDFLTDDLPDRALTIHQFSHF
                    LNSNLYPMLSYSSVILVLRPDV"
BASE COUNT      619 a    546 c    661 g    611 t
ORIGIN


  Query Match           53.8%;  Score 1611.8;  DB 6;  Length 2437;
  Best Local Similarity 77.5%;  Pred. No. 2.1e-292;
  Matches 2293;  Conservative    0;  Mismatches   37;  Indels  629;  Gaps    3;

Qy      4 cggggccggcgccgcgcctgcgacggtggcgacccaatgcgacggcggggaaggggggtcg 63
          || |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db     11 CGAGGCCGGCGCCGCGCCTGCGACGGTGGCGGCCCAATGCGACGGCGGGGAAGGGGGTCG 70

Qy     64 gcgaggtgtgcgccgcggttgtcgaggcggcgacgaaggtagaggacgaggggggaggagg 123
          ||||||| |||||||||||||||||| |||||||||||  ||||| ||||||||||||||
Db     71 GCGAGGTCTGCGCCGCGGTTGTCGAGGTGGCGACGAAGGCCGAGGATGAGGGGGAGGAGG 130

Qy    124 acgagccggtggcggaggacaggtacgcgctcggcggcgcgtgcagggtgctcgccggaa 183
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    131 ACGAGCCGGTGGCGGAGGACAGGTACGCGCTCGGCGGCGCGTGCAGGGTGCTCGCCGGAA 190

Qy    184 tgcccgcgccgctgggcgccaccgcgctcgccggcggggtcaatttcgccgtctattccg 243
          ||||| |||||||||||||||||||||||||||||||||||||||||||||||| |||| 
Db    191 TGCCCACGCCGCTGGGCGCCACCGCGCTCGCCGGCGGGGTCAATTTCGCCGTCTACTCCG 250

Qy    244 gcggagccaccgccgcggcgctctgcctcttcacgccagaagatctcaaggcggtggggt 303
          |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db    251 GCGGAGCCACAGCCGCGGCGCTCTGCCTCTTCACGCCAGAAGATCTCAAGGC-------- 302

Qy    304 tgcctcccgagtagagttcatcagctttgcgtgcgccgcgcgccccttttttgggcctgc 363

Db    303 ------------------------------------------------------------ 302

Qy    364 aatttaagttttgtactggggcaaatgctgcaggatagggtgaccgaggaggttcccctt 423
                                                 |||||||||||| ||||||||||||||
Db    303 --------------------------------GGATAGGGTGACGGAGGAGGTTCCCCTT 330

Qy    424 gacccccctgatgaatcggaccgggaacgtgtggcatgtcttcatcgaaggcgagctgcac 483
          |||||||||||||||||| |||||||| |||||||||||||||||||||||||||||||||
Db    331 GACCCCCTGATGAATCGGACTGGGAACGTATGGCATGTCTTCATCGAAGGCGAGCTGCAG 390

Qy    484 aacatgctttacgggtacaggttcgacggcacctttgctcctcactgcgggcactacctt 543
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    391 GACATGCTTTACGGGTACAGGTTCGACGGCACCTTTGCTCCTCACTGCGGGCACTACCTT 450

Qy    544 gatgtttccaatgtcgtggtggatccttatgctaaggcagtgataagccgaggggagtat 603
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    451 GATGTTTCCAATGTCGTGGTGGATCCTTATGCTAAGGCAGTGATAAGCCGAGGGGAGTAT 510

Qy    604 ggtgttccagcgcgtggtaacaattgctggcctcagatggctggcatgatccctcttcca 663
          |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    511 GGTGTTCCGGCGCGTGGTAACAATTGCTGGCCTCAGATGGCTGGCATGATCCCTCTTCCA 570

Qy    664 tatagcacgtttgattgggaaggcgacctacctctaagatatcctcaaaaggacctggta 723
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    571 TATAGCACGTTTGATTGGGAAGGCGACCTACCTCTAAGATATCCTCAAAAGGACCTGGTA 630

Qy    724 atatatgagatgcacttgcgtggattcacgaagcatgattcaagcaatgtagaacatccg 783
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    631 ATATATGAGATGCACTTGCGTGGATTCACGAAGCATGATTCAAGCAATGTAGAACATCCC 690

Qy    784 ggtactttcattggagctgtgtcgaagcttgactatttgaaggagcttggagttaattgt 843
          |||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db    691 GGTACTTTCATTGGGGCTGTGTCGAAGCTTGACTATTTGAAGGAGCTTGGAGTTAATTGT 750

Qy    844 attgaattaatgccctgccatgagttcaacgagctggagtactcaacctcttcttccaag 903
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    751 ATTGAGTTAATGCCCTGCCATGAGTTCAACGAGCTGGAGTACTCAACCTCTTCTTCCAAG 810

Qy    904 atgaacttttggggatattctaccataaacttcttttcaccaatgacaagatacacatca 963
          |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||

Db    811 ATGAACTTTTGGGGATATTCTACCATAAACTTCTTTTCACCAATGACGAGATACACATCA 870

Qy    964 ggcgggataaaaaactgtgggcgtgatgccataaatgagttcaaaacttttgtaagagag 1023
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    871 GGCGGGATAAAAAACTGTGGGCGTGATGCCATAAATGAGTTCAAAACTTTTGTAAGAGAG 930

Qy   1024 gctcacaaacggggaattgaggtgatcctggatgttgtcttcaaccatacagctgagggt 1083
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    931 GCTCACAAACGGGGAATTGAGGTGATCCTGGATGTTGTCTTCAACCATACAGCTGAGGGT 990

Qy   1084 aatgagaatggtccaatattatcatttaaggggtcgataatactacatactatatgctt 1143
          |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db    991 AATGAGAATGGTCCAATATTATCATTTAGGGGGGTCGATAATACTACATACTATATGCTT 1050

Qy   1144 gcacccaagggagagttttataactattctggctgtgggaataccttcaactgtaatcat 1203
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1051 GCACCCAAGGGAGAGTTTTATAACTATTCTGGCTGTGGGAATACCTTCAACTGTAATCAT 1110

Qy   1204 cctgtggttcgtcaattcattgtagattgtttaagatactgggtgacggaaatgcatgtt 1263
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1111 CCTGTGGTTCGTCAATTCATTGTAGATTGTTTAAGATACTGGGTGACGGAAATGCATGTT 1170

Qy   1264 gatggttttcgttttgatcttgcatccataatgaccagaggttccagtctgtgggatcca 1323
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1171 GATGGTTTTCGTTTTGATCTTGCATCCATAATGACCAGAGGTTCCAGTCTGTGGGATCCA 1230

Qy   1324 gttaacgtgtatggagctccaatagaaggtgacatgatcacaacagggacacctcttgtt 1383
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1231 GTTAACGTGTATGGAGCTCCAATAGAAGGTGACATGATCACAACAGGGACACCTCTTGTT 1290

Qy   1384 actccaccacttattgacatgatcagcaatgacccaattcttggaggcgtcaagctcatt 1443
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db   1291 ACTCCACCACTTATTGACATGATCAGCAATGACCCAATTCTTGGAGGCGTCAAGCTCGTT 1350

Qy   1444 gctgaagcatgggatgcaggaggcctctatcaagtaggtcaattccctcactggaatgtt 1503
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1351 GCTGAAGCATGGGATGCAGGAGGCCTCTATCAAGTAGGTCAATTCCCTCACTGGAATGTT 1410

Qy   1504 tggtctgagtggaatgggaagtaccgggacattgtgcgtcaattcattaaaggcactgat 1563
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1411 TGGTCTGAGTGGAATGGGAAGTACCGGGACATTGTGCGTCAATTCATTAAAGGCACTGAT 1470

Qy   1564 ggatttgctggtggttttgccgaatgtctttgtggaagtccacacctataccaggtaagt 1623
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1471 GGATTTGCTGGTGGTTTTGCCGAATGTCTTTGTGGAAGTCCACACCTATAC--------- 1521

Qy   1624 tgtggcaatacttgtaaatgagttgagtgaatgtcacctggatttttatatataccaca 1683

Db   1522 ------------------------------------------------------------ 1521

Qy   1684 tgatgatacacatctaaatatataacaatcatagtgtatgcatatgcatttggctaagaa 1743

Db   1522 ------------------------------------------------------------ 1521

Qy   1744 gtattagtgtatacactagtgctatatataggttttaacacccaacttgccaatgaagga 1803

Db   1522 ------------------------------------------------------------ 1521

Qy   1804 acatagggctttctagttatcttatttatttgtccggtgaataatccactgaaaaattcc 1863

Db   1522 ------------------------------------------------------------ 1521

Qy   1864 agccatgtcatttttttaggggggagaagaaactatattgatttgccccccctaaaagaag 1923

Db   1522 ------------------------------------------------------------ 1521

Qy   1924 ccatctcagaattcataggtaagttgcttttctgtaaagaaaggaaaacgacttcatact 1983

Db   1522 ------------------------------------------------------------ 1521

Qy   1984 ttctatcggtgctaacttagctcgatgtatatttgtaagatgaatgccaaatttaatttg 2043

Db   1522 ------------------------------------------------------------ 1521
```

```
Qy   2044 tcggataatttgatctgttattcacaaatttctatttggtttctctagaaatcaaaccag 2103

Db   1522 ------------------------------------------------------------ 1521

Qy   2104 taacttgttattggcactgcaacttcttattgattaatcaggcaggaggaaggaaaccctt 2163
                                                 ||||||||||||||||||||||
Db   1522 -------------------------------------CAGGCAGGAGGAAGGAAACCTT 1543

Qy   2164 ggcacagtatcaactttgtatgtgcacatgatggatttacactggctgatttggtaacat 2223
          |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db   1544 GGCACAGTATCAACTTTGTATGTGCACACGATGGATTTACACTGGCTGATTTGGTAACAT 1603

Qy   2224 ataataagaagtacaatttaccaaatggggagaacaacagagatggagaaaatcacaatc 2283
          ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1604 ATAATAACAAGTACAATTTACCAAATGGGGAGAACAACAGAGATGGAGAAAATCACAATC 1663

Qy   2284 ttagctggaattgtggggaggaaggagaattcgcaagattgtctgtcaaaagattgagga 2343
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1664 TTAGCTGGAATTGTGGGGAGGAAGGAGAATTCGCAAGATTGTCTGTCAAAAGATTGAGGA 1723

Qy   2344 agaggcagatgcgcaatttctttgtttgtctcatggtttctcaaggagttccaatgttct 2403
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1724 AGAGGCAGATGCGCAATTTCTTTGTTTGTCTCATGGTTTCTCAAGGAGTTCCAATGTTCT 1783

Qy   2404 acatgggtgatgaatatggccacacaaaaggggggcaacaacaatacatactgccatgatt 2463
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1784 ACATGGGTGATGAATATGGCCACACAAAAGGGGGCAACAACAATACATACTGCCATGATT 1843

Qy   2464 cttatgtcaattatttttcgctgggataaaaaagaacaatactctgagttgcaccgattct 2523
          ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db   1844 CTTATGTCAATTATTTTCGCTGGGATAAAAAAGAACAATACTCTGACTTGCACCGATTCT 1903

Qy   2524 gctgcctcatgaccaaattccgcaaggagtgcgagggtcttggccttgaggactttccaa 2583
          | ||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db   1904 GTTGCCTCATGACCAAATTCCGCAAGGAGTGCGAGGGTCTTGGCCTTGAGGATTTTCCAA 1963

Qy   2584 cggccaaacggctgcagtggcatggtcatcagcctgggaagcctgattggtctgagaata 2643
          |||||` ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1964 CGGCCGAACGGCTGCAGTGGCATGGTCATCAGCCTGGGAAGCCTGATTGGTCTGAGAATA 2023

Qy   2644 gccgattcgttgccttttccatgaaagatgaaagacagggcgagatctatgtggccttca 2703
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2024 GCCGATTCGTTGCCTTTTCCATGAAAGATGAAAGACAGGGCGAGATCTATGTGGCCTTCA 2083

Qy   2704 acaccagccacttaccggccgttgttgagctcccagagcgcgcagggcgccggtgggaac 2763
          |||||||||||||||||||||||||||||| |||||| ||||||||||||||||||||
Db   2084 ACACCAGCCACTTACCGGCCGTTGTTGAGCTCCCGGAGCGCACAGGGCGCCGGTGGGAAC 2143

Qy   2764 cggtggtggacacaggcaagccagcaccatacgacttcctcaccgacgacttacctgatc 2823
          |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db   2144 CGGTGGTGGACACAGGCAAGCCAGCACCATACGACTTCCTCACTGACGACTTACCTGATC 2203

Qy   2824 gcgctctcaccatacaccagttctcgcatttcctctactccaacctctaccccatgctca 2883
          |||||||||||||||||||||||||||| |||||||| ||||||||||||||||||||||
Db   2204 GCGCTCTCACCATACACCAGTTCTCTCATTTCCTCAACTCCAACCTCTACCCCATGCTCA 2263

Qy   2884 gctactcatcggtcatcctagtattgcgccctgatgtttgagagaccaatatatacagta 2943
          |||||||||||||||||||||||||||||||||||||||||| |    ||||||||||
Db   2264 GCTACTCATCGGTCATCCTAGTATTGCGCCCTGATGTTTGAGAGGC--GGATATACAGTA 2321

Qy   2944 aataatatgtctatatgta 2962
          |||||||||| ||||||||
Db   2322 AATAATATGTATATATGTA 2340


RESULT   7
AB074189
LOCUS       AB074189               2584 bp    mRNA    linear   PLN 14-NOV-2001
DEFINITION  Hordeum vulgare hvisol gene for isoamylase, complete cds.
ACCESSION   AB074189
VERSION     AB074189.1  GI:16923278
KEYWORDS    .
SOURCE      Hordeum vulgare (cultivar:Shikoku-Hadaka 84) immature seed cDNA to
            mRNA, clone_lib:Lamda zap II barlry seed cDNA library.
  ORGANISM  Hordeum vulgare
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Hordeum.
REFERENCE   1  (bases 1 to 2584)
  AUTHORS   Komatsu,A. and Komae,K.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-NOV-2001) Akira Komatsu, National Institute of Crop
            Science, Laboratory of Wheat and Barley Quality; 2-1-18 Kannondai,
            Tsukuba, Ibaraki 305-8518, Japan (E-mail:akomatsu@naro.affrc.go.jp,
            URL:http://nics.naro.affrc.go.jp/, Tel:81-298-38-8949,
            Fax:81-298-38-8949)
FEATURES             Location/Qualifiers
     source          1. .2584
                     /organism="Hordeum vulgare"
                     /cultivar="Shikoku-Hadaka 84"
                     /db_xref="taxon:4513"
                     /clone_lib="Lamda zap II barlry seed cDNA library"
                     /dev_stage="immature seed"
                     /country="Japan:Ibaraki,Tsukuba"
     transit_peptide 103. .246
                     /gene="hvisol"
     CDS             103. .2472
                     /gene="hvisol"
                     /codon_start=1
                     /product="isoamylase"
                     /protein_id="BAB72000.1"
                     /db_xref="GI:16923279"
                     /translation="MKMMAMAEAPCLCARPFLAARAWRPGPAPRLRRWRPNATAGKGV
                     GEVCAAAVEAATKAEEEEDEEEEPVAEDRYALGGACRVLAGMPAPLGATALAGGVNFA
                     VYSGGATAAALCLFTPEDLKADRVSEEVPLDPLMNRTGDVWHVFLEGELHGMLYGYRF
                     DGTFAPHCGHYFDVSNVVVDPYAKAVISREEYGVPAHGNNCWPQMAGMIPLPYSTFDW
                     EGDLPLRYPQKDLVIYEMHLRGFTKHDSSNVEHPGTFIGAVSKLDYLKELGVNCIELM
                     PCHEFNELEYATSSSKMNFWGYSTINFFSPMTRYTSGGIKNCGRDGINEFKTFVRESH
                     KRGIEVILDVVFNHTAEGNENGPILSFRGVDNTTYYMLAPKGEFYNYSGCGNTFNCNH
                     PVVRQFIVDCLRYWVMEMHIDGFRFDLASIMTRGSSLWDPINVYGAPIEGDMITTGTP
                     LVTPPLIDMISNDPILGGVKLIAEAWDAGGLYQVGQFPHWNVWSEWNGKYRDIVRQFI
                     KGTDGFAGGFAECLCGSPQLYQAGGRKPWHSINFVCAHDGFTLADLVTYNTKYNLPNG
                     EDIRDGENHNLSWNCGEEGEFARSSVKRLRKRQMRNFFVCLMVSQGVPMFYMGDEYGH
                     TKGGNNNTYCHDSYVNYFRWDKKEEHSDLHRFCCLMTKFRKECEGLGLEDFPTAEQLQ
                     WHGHQPGKPDWSEKSRFVAFSMKDETKGEIYVAFNTSHLAAVVELPERIGHRWEPVVD
                     TGKPAPYDFLTDDLPDRALTVQQFSHFLNSNLYPMLSYTSVILVSRPDV"
     gene            103. .2472
                     /gene="hvisol"
     mat_peptide     247. .2469
                     /gene="hvisol"
                     /product="unnamed"
BASE COUNT      656 a    606 c    701 g    621 t
ORIGIN


  Query Match           50.4%;  Score 1510.8;  DB 8;  Length 2584;
  Best Local Similarity   75.6%;  Pred. No. 2e-273;
  Matches 2238;  Conservative    0;  Mismatches   92;  Indels  632;  Gaps    4;

Qy       1 ggtcggggccggcgccgcgcctgcgacggtggcgacccaatgcgacggcggggaaggggg 60
           || |||||||||||||||||||||| ||||| ||||||||||||||||||||||||||
Db     173 GGCCGGGGCCGGCGCCGCGCCTGCGACGATGGCGGCCCAATGCGACGGCGGGGAAGGGGG 232

Qy      61 tcggcgaggtgtgcgccgcggttgtcgaggcggcgacgaaggtagaggacgagggg---g 117
           |||||||||| |||||||||  ||||||||||||||||||  ||||| |||| |   |
Db     233 TCGGCGAGGTCTGCGCCGCGGCCGTCGAGGCGGCGACGAAGGCCGAGGAGGAGGAGGACG 292

Qy     118 aggaggacgagccggtggcggaggacaggtacgcgctcggcggcgcgtgcagggtgctcg 177
           ||||||| |||||||||| |||||||||| |||||||| ||||||||||||||||||||
Db     293 AGGAGGAGGAGCCGGTGGCTGAGGACAGGTATGCGCTCGGAGGCGCGTGCAGGGTGCTCG 352

Qy     178 ccggaatgcccgcgccgctgggcgccaccgcgctcgccggcggggtcaatttcgccgtct 237
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     353 CCGGAATGCCCGCGCCGCTGGGCGCCACCGCGCTCGCCGGCGGGGTCAATTTCGCCGTCT 412
```

```
Qy    238 attccggcggagccaccgccgcggcgctctgcctcttcacgccagaagatctcaaggcgg 297
          | || |||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db    413 ACTCAGGCGGAGCCACCGCCGCTGCGCTCTGCCTCTTCACGCCAGAAGATCTCAAGGC-- 470

Qy    298 tggggttgcctcccgagtagagttcatcagctttgcgtgcgccgcgcgccccttttttgg 357

Db    471 ------------------------------------------------------------ 470

Qy    358 gcctgcaatttaagttttgtactggggcaaatgctgcaggataggg tgaccgaggaggtt 417
                                                      |||||||||| ||||||||||
Db    471 -----------------------------------GGATAGGGTGAGCGAGGAGGTT 492

Qy    418 ccccttgacccccctgatgaatcggaccgggaacgtgtggcatgtcttcatcgaaggcgag 477
          |||||||||||||||||||||||| ||| |||||||||||||||| ||||||||||
Db    493 CCCCTTGACCCCCTGATGAATCGGACTGGGGACGTGTGGCATGTCTTCCTCGAAGGCGAG 552

Qy    478 ctgcacaacatgctttacgggtacaggttcgacggcacctttgctcctcactgcgggcac 537
          ||||||  |||||||| |||||||||||||||||||| |||||||||||||||||||
Db    553 CTGCACGGCATGCTTTATGGGTACAGGTTCGACGGCACCTTCGCTCCTCACTGCGGGCAC 612

Qy    538 taccttgatgtttccaatgtcgtggtggatccttatgctaaggcagtgataagccgaggg 597
          ||| |||||||||||||| ||||||||||||||||||||||||||||||||||| | |
Db    613 TACTTTGATGTTTCCAATGTTGTGGTGGATCCTTATGCTAAGGCAGTGATAAGCCGGGAG 672

Qy    598 gagtatggtgttccagcgcgtggtaacaattgctggcctcagatggctggcatgatccct 657
          |||||||||||||| |||| ||||||||||||||||||||||||||||||||||||||
Db    673 GAGTATGGTGTTCCGGCGCATGGTAACAATTGCTGGCCTCAGATGGCTGGCATGATCCCT 732

Qy    658 cttccatatagcacgtttgattgggaaggcgacctacctctaagatatcctcaaaaggac 717
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    733 CTTCCATATAGCACGTTTGATTGGGAAGGCGACCTACCTCTAAGATATCCTCAAAAGGAC 792

Qy    718 ctggtaatatatgagatgcacttgcgtggattcacgaagcatgattcaagcaatgtagaa 777
          || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    793 CTTGTAATATATGAGATGCACTTGCGTGGATTCACGAAGCATGATTCAAGCAATGTAGAA 852

Qy    778 catccgggtactttcattggagctgtgtcgaagcttgactatttgaaggagcttggagtt 837
          ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db    853 CATCCGGGTACTTTCATTGGGGCTGTGTCGAAGCTTGACTATTTGAAGGAGCTTGGAGTT 912

Qy    838 aattgtattgaattaatgccctgccatgagttcaacgagctggagtactcaacctcttct 897
          |||||||| ||||||||||||||||||||||||||||||||  ||||||||||
Db    913 AATTGTATAGAATTAATGCCCTGCCATGAGTTCAACGAGCTGGAGTATGCAACCTCTTCT 972

Qy    898 tccaagatgaacttttggggatattctaccataaacttcttttcaccaatgacaagatac 957
          ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db    973 TCCAAGATGAACTTTTGGGGATATTCTACCATAAACTTCTTTTCACCAATGACGAGATAC 1032

Qy    958 acatcaggcgggataaaaaactgtgggcgtgatgccataaatgagttcaaaacttttgta 1017
          || ||||| |||||||||||||||||||||||||| |||||| ||||||||||||||||||
Db   1033 ACGTCAGGTGGGATAAAAAACTGTGGGCGTGATGGCATAAACGAGTTCAAAACTTTTGTA 1092

Qy   1018 agagaggctcacaaacggggaattgaggtgatcctggatgttgtcttcaaccatacagct 1077
          |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1093 AGAGAGTCTCACAAACGGGGAATTGAGGTGATCCTGGATGTTGTCTTCAACCATACAGCT 1152

Qy   1078 gagggtaatgagaatggtccaatattatcatttaagggggtcgataatactacatactat 1137
          |||||||||||||||||| |||||||||||| |||||| ||||||||||||||||||||
Db   1153 GAGGGTAATGAGAATGGTCCGATATTATCATTTAGGGGGGTTGATAATACTACATACTAT 1212

Qy   1138 atgcttgcacccaagggagagttttataactattctggctgtgggaataccttcaactgt 1197
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1213 ATGCTTGCACCCAAGGGAGAGTTTTATAACTATTCTGGCTGTGGGAATACCTTCAACTGT 1272

Qy   1198 aatcatcctgtggttcgtcaattcattgtagattgtttaagatactgggtgacggaaatg 1257
          |||||||||||||||||||||| |||||||||||||||||||||||| |||||||
Db   1273 AATCATCCTGTGGTTCGTCAATTTATTGTAGATTGTTTAAGATACTGGGTGATGGAAATG 1332

Qy   1258 catgttgatggttttcgttttgatcttgcatccataatgaccagaggttccagtctgtgg 1317
          ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1333 CATATTGATGGTTTTCGTTTTGATCTTGCATCCATAATGACCAGAGGTTCCAGTCTGTGG 1392

Qy   1318 gatccagttaacgtgtatggagctccaatagaaggtgacatgatcacaacagggacacct 1377
          |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1393 GATCCAATTAACGTGTATGGAGCTCCAATAGAAGGTGACATGATCACAACAGGGACACCT 1452

Qy   1378 cttgttactccaccacttattgacatgatcagcaatgacccaattcttggaggcgtcaag 1437
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1453 CTTGTTACTCCACCACTTATTGACATGATCAGCAATGACCCAATTCTTGGAGGCGTCAAG 1512

Qy   1438 ctcattgctgaagcatgggatgcaggaggcctctatcaagtaggtcaattccctcactgg 1497
          |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db   1513 CTCATTGCTGAAGCATGGGATGCAGGAGGCCTCTATCAAGTAGGTCAGTTCCCTCACTGG 1572

Qy   1498 aatgtttggtctgagtggaatgggaagtaccgggacattgtgcgtcaattcattaaaggc 1557
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1573 AATGTTTGGTCTGAGTGGAATGGGAAGTACCGGGACATTGTGCGTCAATTCATTAAAGGC 1632

Qy   1558 actgatggatttgctggtggttttgccgaatgtctttgtggaagtccacacctataccag 1617
          |||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db   1633 ACTGATGGATTTGCTGGTGGTTTTGCCGAATGTCTTTGTGGAAGTCCACAACTATAC--- 1689

Qy   1618 gtaagttgtggcaatacttgtaaatgagttgagtgaatgtcacctggatttttatatat 1677

Db   1690 ------------------------------------------------------------ 1689

Qy   1678 accacatgatgatacacatctaaatatataacaatcatagtgtatgcatatgcatttggc 1737

Db   1690 ------------------------------------------------------------ 1689

Qy   1738 taagaagtattagtgtatacactagtgctatatataggttttaacacccaacttgccaat 1797

Db   1690 ------------------------------------------------------------ 1689

Qy   1798 gaaggaacatagggctttctagttatcttatttatttgtccggtgaataatccactgaaa 1857

Db   1690 ------------------------------------------------------------ 1689

Qy   1858 aattccagccatgtcatttttttaggggggggagaagaaactatattgatttgccccccta 1917

Db   1690 ------------------------------------------------------------ 1689

Qy   1918 aagaagccatctcagaattcataggtaagttgcttttctgtaaagaaaggaaaacgactt 1977

Db   1690 ------------------------------------------------------------ 1689

Qy   1978 catactttctatcggtgctaacttagctcgatgtatatttgtaagatgaatgccaaattt 2037

Db   1690 ------------------------------------------------------------ 1689

Qy   2038 aatttgtcggataatttgatctgttattcacaaatttctatttggtttctctagaaatca 2097

Db   1690 ------------------------------------------------------------ 1689

Qy   2098 aaccagtaacttgttattggcactgcaacttcttattgattaatcaggcaggaggaagga 2157
                                                      ||||||||||||||||
Db   1690 ---------------------------------------------CAGGCAGGAGGAAGGA 1705

Qy   2158 aaccttggcacagtatcaactttgtatgtgcacatgatggatttacactggctgatttgg 2217
          |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db   1706 AACCTTGGCACAGTATCAATTTTGTATGTGCACATGATGGATTTACACTGGCTGATTTGG 1765

Qy   2218 taacatataataagaagtacaatttaccaaatggggagaacaacagagatggagaaaatc 2277
          |||||||||||| |||||||||||||||||||||| ||| ||||||||||||||||||
Db   1766 TAACATATAATACCAAGTACAATTTACCAAATGGGGAGGACATCAGAGATGGAGAAAATC 1825

Qy   2278 acaatcttagctggaattgtggggaggaaggagaattcgcaagattgtctgtcaaaagat 2337
          |||||||||||||||| |||||||||||||||||||||||||||| ||| ||||||||||
Db   1826 ACAATCTTAGCTGGAACTGTGGGGAGGAAGGAGAATTCGCAAGATCGTCAGTCAAAAGAT 1885

Qy   2338 tgaggaagaggcagatgcgcaatttctttgtttgtctcatggtttctcaaggagttccaa 2397
          |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db   1886 TGAGGAAGAGGCAGATGCGCAATTTCTTTGTCTGTCTCATGGTTTCTCAAGGAGTTCCAA 1945

Qy   2398 tgttctacatgggtgatgaatatggccacacaaaaggggcaacaacaatacatactgcc 2457
          |||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
```

```
Db    1946 TGTTCTACATGGGCGATGAATATGGCCACACAAAAGGGGGCAACAACAATACATACTGCC 2005

Qy    2458 atgattcttatgtcaattatttttcgctgggataaaaaaagaacaatactctgagttgcacc 2517
           |||||||||||||||||||||||||||||||||||||||||||||||| || ||||||| |||||||
Db    2006 ATGATTCTTATGTCAATTATTTTCGCTGGGATAAAAAAGAAGAACACTCTGACTTGCACC 2065

Qy    2518 gattctgctgcctcatgaccaaattccgcaaggagtgcgagggtcttggccttgaggact 2577
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2066 GATTCTGCTGCCTCATGACCAAATTCCGCAAGGAGTGCGAGGGTCTTGGCCTTGAGGACT 2125

Qy    2578 ttccaacggccaaacggctgcagtggcatggtcatcagcctgggaagcctgattggtctg 2637
           |||||||||| |||  ||||| ||||||||||||||||||||||||||||||||||||||||||||
Db    2126 TTCCAACGGCCGAACAACTGCAATGGCATGGTCATCAGCCTGGGAAGCCTGATTGGTCTG 2185

Qy    2638 agaatagccgattcgttgccttttccatgaaagatgaaagacagggcgagatctatgtgg 2697
           |||| |||||||| |||||||||| |||||||||||||| | | ||||||||||||||||||
Db    2186 AGAAAAGCCGATTTGTTGCCTTTTCAATGAAAGATGAAACAAAAGGCGAGATCTATGTGG 2245

Qy    2698 ccttcaacaccagccacttaccggccgttgttgagctcccagagcgcgcagggcgccggt 2757
           |||||||||||||||||||| |||||||||||| |||||||||||  ||||| |||||
Db    2246 CCTTCAACACCAGCCACTTAGCGGCCGTTGTTGAACTCCCAGAGCGCATAGGGCACCGGT 2305

Qy    2758 gggaaccggtggtggacacaggcaagccagcaccatacgacttcctcaccgacgacttac 2817
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db    2306 GGGAACCGGTGGTGGACACAGGCAAGCCAGCACCATACGACTTCCTCACCGACGACCTAC 2365

Qy    2818 ctgatcgcgctctcaccatacaccagttctcgcatttcctctactccaacctctacccca 2877
           |||||||||||||||| |||| ||||| || |||||||| ||||||||||||||||||||||
Db    2366 CTGATCGCGCTCTCACCGTACAGCAGTTTTCTCATTTCCTCAACTCCAACCTCTACCCCA 2425

Qy    2878 tgctcagctactcatcggtcatcctagtattgcgccctgatgtttgagagaccaatatat 2937
           |||||||||| |||||||||||| |||| ||||||||||||||||||||||  || |
Db    2426 TGCTCAGCTACACATCGGTCATCCTTGTATCACGCCCTGATGTTTGAGAGACCA--ATGT 2483

Qy    2938 acagtaaataatatgtctatat 2959
           |||||||||||||||| ||| |
Db    2484 ACAGTAAATAATATGTATATGT 2505


RESULT   8
AF142589
LOCUS       AF142589                1708 bp    mRNA    linear   PLN 12-AUG-1999
DEFINITION  Hordeum vulgare isoamylase 1 mRNA, partial cds.
ACCESSION   AF142589
VERSION     AF142589.1  GI:4929174
KEYWORDS    .
SOURCE      barley.
  ORGANISM  Hordeum vulgare
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Hordeum.
REFERENCE   1  (bases 1 to 1708)
  AUTHORS   Sun,C., Sathish,P., Ahlandsberg,S. and Jansson,C.
  TITLE     Analyses of isoamylase gene activity in wild-type barley indicate
            its involvement in starch synthesis
  JOURNAL   Plant Mol. Biol. 40 (3), 431-443 (1999)
  MEDLINE   99364538
   PUBMED   10437827
REFERENCE   2  (bases 1 to 1708)
  AUTHORS   Sun,C., Sathish,P., Ahlandsberg,S. and Jansson,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-APR-1999) Plant Biology, Genetic Center, The Swedish
            University of Agricultural Sciences, P.O. Box 7080, Uppsala
            S-75007, Sweden
FEATURES             Location/Qualifiers
     source          1. .1708
                     /organism="Hordeum vulgare"
                     /cultivar="Bomi"
                     /db_xref="taxon:4513"
     CDS             <1. .>1708
                     /codon_start=1
                     /product="isoamylase 1"
                     /protein_id="AAD33889.1"
                     /db_xref="GI:4929175"
                     /translation="LGGACRVLAGMPAAGATALAGGVNFAVYSAEPPRALCLFTPEDL
                     KADRVSEEVPLDPLMNRTGDVWHVFLEGELHGMLYGYRFDGTFAPHCGHYFDVSNVVV
                     DPYAKAVISREEYGVPAHGNNSWPQMAGMIPLPYSTFDWEGDLPLRYPQKDLVIYEMH
                     LRGFTKHDSSNVEHPGTFIGAVSKLDYLKELGVNCIELMPCHEFNELEYATSSSKMNF
                     WGYSTINFFSPMTRYTSGGIKNCGRDGINEFKTFVRESHKRGIEVILDVVFNHTAEGN
                     ENGPILSFRGVDNTTYYMLAPKGEFYNYSGCGNTFNCNHPVVRQFIVDCLRYWVMEMH
                     IDGFRFDLASIMTRGSSLWDPINVYGAPIEGDMITTGTPLVTPPLIDMISNDPILGGV
                     KLIAEAWDAGGLYQVGQFPHWNVWSEWNGKYRDIVRQFIKGTDGFAGGFAECLCGSPQ
                     LYQAGGRKPWHSINFVCAHDGFTLADLVTYHNKYNLPNGEDNRDGENHNLSWNCGEEG
                     EFARSSVKRLRKRQMRNFFVCLMVSQGVPIFYMGDEYGHTKGGNNNTYCHDSYVNYFR
                     WEK"
BASE COUNT      444 a    353 c    450 g    461 t
ORIGIN
```

Query Match 39.1%; Score 1171.6; DB 8; Length 1708;
Best Local Similarity 96.5%; Pred. No. 8.4e-210;
Matches 1198; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

```
Qy     396 ggataggg tgaccgaggaggttcccctt gacccccgatgaatcggaccgggaacgtgtg 455
           |||||||||| |||||||||||||||||||||||||||||||||||||||||| ||| |||||||
Db     138 GGATAGGGTGAGCGAGGAGGTTCCCCTTGACCCCCTGATGAATCGGACTGGGGACGTGTG 197

Qy     456 gcatgtcttcatcgaaggcgagctgcacaacatgctttacgggtacaggttcgacggcac 515
           |||||||||| |||||||||||||||||  |||||||| ||||||||||||||||||||||
Db     198 GCATGTCTTCCTCGAAGGCGAGCTGCACGGCATGCTTTATGGGTACAGGTTCGACGGCAC 257

Qy     516 ctttgctcctcactgcgggcactaccttgatgtttccaatgtcgtggtggatccttatgc 575
           ||| |||||||||||||||||||| |||||||||||||||| ||||||||||||||||||
Db     258 CTTCGCTCCTCACTGCGGGCACTACTTTGATGTTTCCAATGTTGTGGTGGATCCTTATGC 317

Qy     576 taaggcagtgataagccgaggggagtatggtgttccagcgcgtggtaacaattgctggcc 635
           |||||||||||||||||| | |||||||||||||||| |||| |||||||||| ||||||
Db     318 TAAGGCAGTGATAAGCCGGGAGGAGTATGGTGTTCCGGCGCATGGTAACAATTCCTGGCC 377

Qy     636 tcagatggctggcatgatccctcttccatatagcacgtttgattgggaaggcgacctacc 695
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     378 TCAGATGGCTGGCATGATCCCTCTTCCATATAGCACGTTTGATTGGGAAGGCGACCTACC 437

Qy     696 tctaagatatcctcaaaaggacctggtaatatatgagatgcacttgcgtggattcacgaa 755
           |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db     438 TCTAAGATATCCTCAAAAGGACCTTGTAATATATGAGATGCACTTGCGTGGATTCACGAA 497

Qy     756 gcatgattcaagcaatgtagaacatccgggtactttcattggagctgtgtcgaagcttga 815
           |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db     498 GCATGATTCAAGCAATGTAGAACATCCGGGTACTTTCATTGGGGCTGTGTCGAAGCTTGA 557

Qy     816 ctatttgaaggagcttggagttaattgtattgaattaatgccctgccatgagttcaacga 875
           |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db     558 CTATTTGAAGGAGCTTGGAGTTAATTGTATAGAATTAATGCCCTGCCATGAGTTCAACGA 617

Qy     876 gctggagtactcaacctcttcttccaagatgaacttttggggatattctaccataaacttt 935
           |||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     618 GCTGGAGTATGCAACCTCTTCTTCCAAGATGAACTTTTGGGGATATTCTACCATAAACTT 677

Qy     936 cttttcaccaatgacaagatacacatcaggcgggataaaaaactgtgggcgtgatgccat 995
           ||||||||||||||| |||||||| ||||| |||||||||||||||||||||||||| |||
Db     678 CTTTTCACCAATGACGAGATACACGTCAGGTGGGATAAAAAACTGTGGGCGTGATGGCAT 737

Qy     996 aaatgagttcaaaacttttgtaagagaggctcacaaacggggaattgaggtgatcctgga 1055
           ||| |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db     738 AAACGAGTTCAAAACTTTTGTAAGAGAGTCTCACAAACGGGGAATTGAGGTGATCCTGGA 797

Qy    1056 tgttgtcttcaaccatacagctgagggtaatgagaatggtccaatattatcatttaaggg 1115
           |||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||| |||
Db     798 TGTTGTCTTCAACCATACAGCTGAGGGTAATGAGAATGGTCCGATATTATCATTTAGGGG 857

Qy    1116 ggtcgataatactacatactatatgcttgcacccaagggagagttttataactattctgg 1175
           ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     858 GGTTGATAATACTACATACTATATGCTTGCACCCAAGGGAGAGTTTTATAACTATTCTGG 917

Qy    1176 ctgtgggaataccttcaactgtaatcatcctgtggttcgtcaattcattgtagattgttt 1235
```

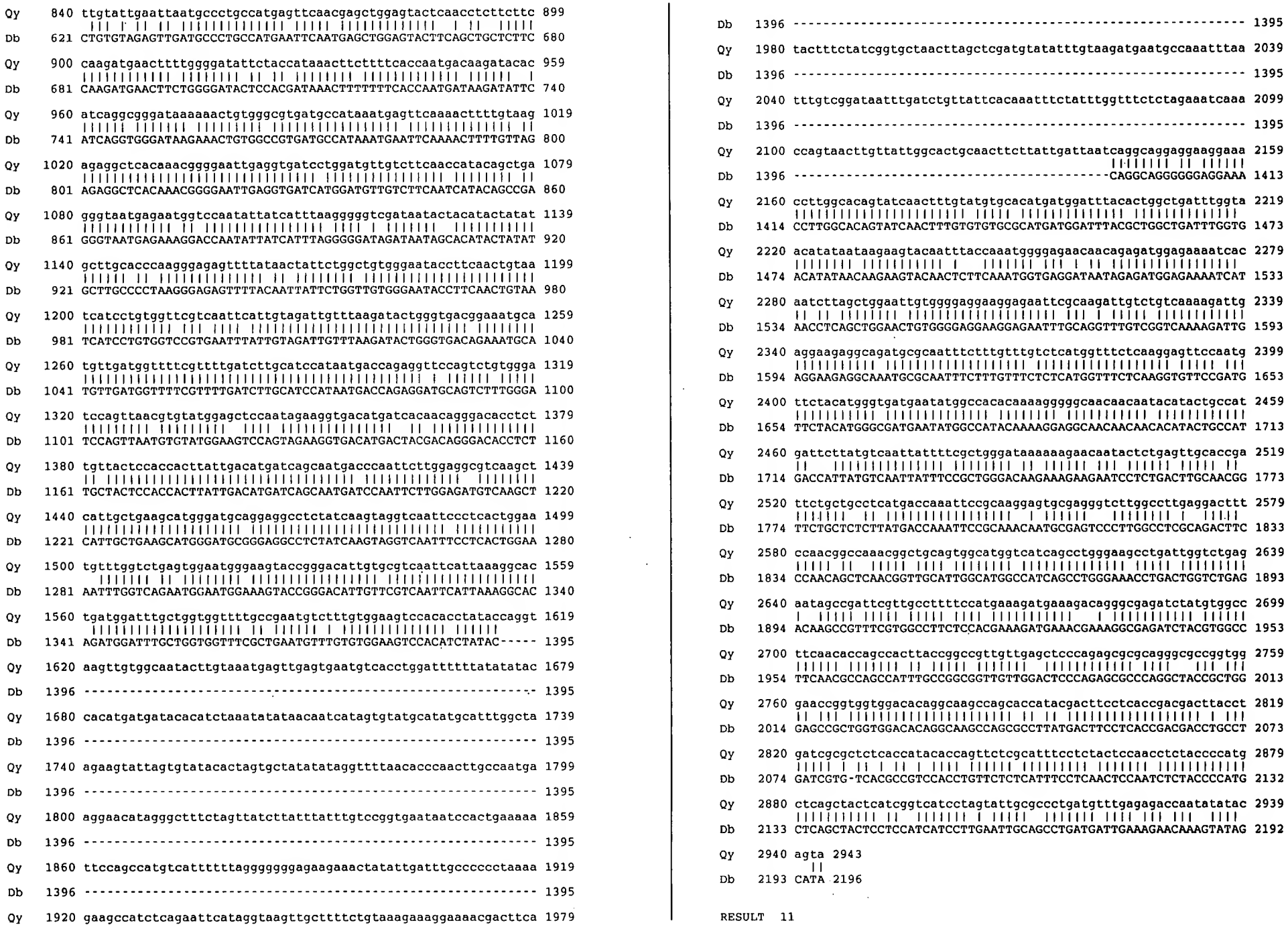
```
Db    918 CTGTGGGAATACCTTCAACTGTAATCATCCTGTGGTTCGTCAATTTATTGTAGATTGTTT 977

Qy   1236 aagatactgggtgacggaaatgcatgttgatggtttttcgttttgatcttgcatccataat 1295
          ||||||||||||| |||||||||| ||||||||||||||||||||||||||||||||||||
Db    978 AAGATACTGGGTGATGGAAATGCATATTGATGGTTTTCGTTTTGATCTTGCATCCATAAT 1037

Qy   1296 gaccagaggttccagtctgtgggatccagttaacgtgtatggagctccaatagaaggtga 1355
          |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db   1038 GACCAGAGGTTCCAGTCTGTGGGATCCAATTAACGTGTATGGAGCTCCAATAGAAGGTGA 1097

Qy   1356 catgatcacaacagggacacctcttgttactccaccacttattgacatgatcagcaatga 1415
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1098 CATGATCACAACAGGGACACCTCTTGTTACTCCACCACTTATTGACATGATCAGCAATGA 1157

Qy   1416 cccaattcttggaggcgtcaagctcattgctgaagcatgggatgcaggaggcctctatca 1475
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1158 CCCAATTCTTGGAGGCGTCAAGCTCATTGCTGAAGCATGGGATGCAGGAGGCCTCTATCA 1217

Qy   1476 agtaggtcaattccctcactggaatgtttggtctgagtggaatgggaagtaccgggacat 1535
          |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1218 AGTAGGTCAGTTCCCTCACTGGAATGTTTGGTCTGAGTGGAATGGGAAGTACCGGGACAT 1277

Qy   1536 tgtgcgtcaattcattaaaggcactgatggatttgctggtggttttgccgaatgtctttg 1595
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1278 TGTGCGTCAATTCATTAAAGGCACTGATGGATTTGCTGGTGGTTTTGCCGAATGTCTTTG 1337

Qy   1596 tggaagtccacacctataccaggtaagttgtggcaatacttg 1637
          |||||||||||| |||||||||| | |  |  | ||  ||||
Db   1338 TGGAAGTCCACAACTATACCAGGCAGGAGGAAGGAAACCTTG 1379
```

```
RESULT   9
AF438329
LOCUS       AF438329               1517 bp    mRNA    linear   PLN 13-NOV-2001
DEFINITION  Triticum aestivum isoamylase precursor, mRNA, complete cds;
            alternatively spliced.
ACCESSION   AF438329
VERSION     AF438329.1  GI:16905064
KEYWORDS    .
SOURCE      bread wheat.
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Triticum.
REFERENCE   1  (bases 1 to 1517)
  AUTHORS   Netrphan,S., Khandelwal,R.L. and Chibbar,R.N.
  TITLE     Characterization of isoamylase transcripts produced during wheat
            grain development and germination
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1517)
  AUTHORS   Netrphan,S., Khandelwal,R.L. and Chibbar,R.N.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-OCT-2001) Plant Biotechnology Institute, 110
            Gymnasium Place, Saskatoon, SK S7N 0W9, Canada
FEATURES             Location/Qualifiers
     source          1. .1517
                     /organism="Triticum aestivum"
                     /cultivar="Fielder"
                     /db_xref="taxon:4565"
                     /tissue_type="endosperm"
     CDS             102. .1424
                     /EC_number="3.2.1.68"
                     /function="hydrolyzes alpha-1,6-glucosidic bonds of
                     starch, amylopectin, and glycogen"
                     /note="alternatively spliced; glycogen 6-glucanohydrolase"
                     /codon_start=1
                     /product="isoamylase precursor"
                     /protein_id="AAL31016.1"
                     /db_xref="GI:16905065"
                     /translation="MTMMAMPEAPCLCARPSLAARARRPGSGPAPRLRRWRPNATAGK
                     GVGEVCAAVVEAATKVEDEGEEDEPVAEDRYALGGACRVLAGMPAPLGATALAGGVNF
                     AVYSGGATAAALCLFTPEDLKADRVTEEVPLDPLMNRTGNVWHVFIEGELHNMLYGYR
```

```
                     FDGTFAPHCGHYLDVSNVVVDPYAKAVISRGEYGVPARGNNCWPQMAGMIPLPYSTFD
                     WEGDLPLRYPQKDLVIYEMHLRGFTKHDSSNVEHPGTFIGAVSKLDYLKELGVNCIEL
                     MPCHEFNELEYSTSSSKMNFWGYSTINFFSPMTRYTSGGIKNCGRDAINEFKTFVREA
                     HKRGIEVILDVVFNHTAEGNENGPILSFRGVDNTTYYMLAPKGEFYNYSGCGNTFNCN
                     HPVVRQFIVDCLRYWVTEMHVDGFRFDLASIMTRGSSLWDPVNVYGAP"
BASE COUNT      361 a    369 c    433 g    354 t
ORIGIN
```

```
  Query Match             37.8%;  Score 1134.2;  DB 8;  Length 1517;
  Best Local Similarity   92.4%;  Pred. No. 8.6e-203;
  Matches 1246;  Conservative    0;  Mismatches    3;  Indels  100;  Gaps    1;

Qy      1 ggtcggggccggcgccgcgcctgcgacggtggcgacccaatgcgacggcggggaaggggg 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    178 GGTCGGGGCCGGCGCCGCGCCTGCGACGGTGGCGACCCAATGCGACGGCGGGGAAGGGGG 237

Qy     61 tcggcgaggtgtgcgccgcggttgtcgaggcggcgacgaaggtagaggacgaggggggagg 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    238 TCGGCGAGGTGTGCGCCGCGGTTGTCGAGGCGGCGACGAAGGTAGAGGACGAGGGGGAGG 297

Qy    121 aggacgagccggtggcggaggacaggtacgcgctcggcggcgcgtgcagggtgctcgccg 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    298 AGGACGAGCCGGTGGCGGAGGACAGGTACGCGCTCGGCGGCGCGTGCAGGGTGCTCGCCG 357

Qy    181 gaatgcccgcgccgctgggcgccaccgcgctcgccggcggggtcaatttcgccgtctatt 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    358 GAATGCCCGCGCCGCTGGGCGCCACCGCGCTCGCCGGCGGGGTCAATTTCGCCGTCTATT 417

Qy    241 ccggcggagccaccgccgcggcgctctgcctcttcacgccagaagatctcaaggcggtgg 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    418 CCGGCGGAGCCACCGCCGCGGCGCTCTGCCTCTTCACGCCAGAAGATCTCAAGGC----- 472

Qy    301 ggttgcctcccgagtagagttcatcagctttgcgtgcgccgcgcgccccttttttgggcc 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    473 ---------------------------------------------------------- 472

Qy    361 tgcaatttaagttttgtactggggcaaatgctgcaggatagggtgaccgaggaggttccc 420
                                                  ||||||||||||||||||||||||
Db    473 ----------------------------------GGATAGGGTGACCGAGGAGGTTCCC 497

Qy    421 cttgacccccctgatgaatcggaccgggaacgtgtggcatgtcttcatcgaaggcgagctg 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    498 CTTGACCCCCTGATGAATCGGACCGGGAACGTGTGGCATGTCTTCATCGAAGGCGAGCTG 557

Qy    481 cacaacatgctttacgggtacaggttcgacggcacctttgctcctcactgcgggcactac 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    558 CACAACATGCTTTACGGGTACAGGTTCGACGGCACCTTTGCTCCTCACTGCGGGCACTAC 617

Qy    541 cttgatgtttccaatgtcgtggtggatccttatgctaaggcagtgataagccgaggggag 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    618 CTTGATGTTTCCAATGTCGTGGTGGATCCTTATGCTAAGGCAGTGATAAGCCGAGGGGAG 677

Qy    601 tatggtgttccagcgcgtggtaacaattgctggcctcagatggctggcatgatccctctt 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    678 TATGGTGTTCCAGCGCGTGGTAACAATTGCTGGCCTCAGATGGCTGGCATGATCCCTCTT 737

Qy    661 ccatatagcacgtttgattgggaaggcgacctacctctaagatatcctcaaaaggacctg 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    738 CCATATAGCACGTTTGATTGGGAAGGCGACCTACCTCTAAGATATCCTCAAAAGGACCTG 797

Qy    721 gtaatatatgagatgcacttgcgtggattcacgaagcatgattcaagcaatgtagaacat 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    798 GTAATATATGAGATGCACTTGCGTGGATTCACGAAGCATGATTCAAGCAATGTAGAACAT 857

Qy    781 ccgggtactttcattggagctgtgtcgaagcttgactatttgaaggagcttggagttaat 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    858 CCGGGTACTTTCATTGGAGCTGTGTCGAAGCTTGACTATTTGAAGGAGCTTGGAGTTAAT 917

Qy    841 tgtattgaattaatgccctgccatgagttcaacgagctggagtactcaacctcttcttcc 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    918 TGTATTGAATTAATGCCCTGCCATGAGTTCAACGAGCTGGAGTACTCAACCTCTTCTTCC 977
```

```
Qy     901 aagatgaacttttggggatattctaccataaacttcttttcaccaatgacaagatacaca 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     978 AAGATGAACTTTTGGGGATATTCTACCATAAACTTCTTTTCACCAATGACAAGATACACA 1037

Qy     961 tcaggcgggataaaaaactgtgggcgtgatgccataaatgagttcaaaacttttgtaaga 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1038 TCAGGCGGGATAAAAAACTGTGGGCGTGATGCCATAAATGAGTTCAAAACTTTTGTAAGA 1097

Qy    1021 gaggctcacaaacggggaattgaggtgatcctggatgttgtcttcaaccatacagctgag 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1098 GAGGCTCACAAACGGGGAATTGAGGTGATCCTGGATGTTGTCTTCAACCATACAGCTGAG 1157

Qy    1081 ggtaatgagaatggtccaatattatcatttaaggggggtcgataatactacatactatatg 1140
           |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db    1158 GGTAATGAGAATGGTCCAATATTATCATTTAGGGGGGTCGATAATACTACATACTATATG 1217

Qy    1141 cttgcacccaagggagagttttataactattctggctgtgggaatacctttcaactgtaat 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1218 CTTGCACCCAAGGGAGAGTTTTATAACTATTCTGGCTGTGGGAATACCTTCAACTGTAAT 1277

Qy    1201 catcctgtggttcgtcaattcattgtagattgtttaagatactgggtgacggaaatgcat 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1278 CATCCTGTGGTTCGTCAATTCATTGTAGATTGTTTAAGATACTGGGTGACGGAAATGCAT 1337

Qy    1261 gttgatggttttcgttttgatcttgcatccataatgaccagaggttccagtctgtgggat 1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1338 GTTGATGGTTTTCGTTTTGATCTTGCATCCATAATGACCAGAGGTTCCAGTCTGTGGGAT 1397

Qy    1321 ccagttaacgtgtatggagctccaataga 1349
           |||||||||||||||||||||||  |||
Db    1398 CCAGTTAACGTGTATGGAGCTCCATGAGA 1426


RESULT  10
AB015615
LOCUS       AB015615               2288 bp    mRNA    linear   PLN 06-JUL-2001
DEFINITION  Oryza sativa mRNA for isoamylase, partial cds.
ACCESSION   AB015615
VERSION     AB015615.1  GI:3252793
KEYWORDS    isoamylase.
SOURCE      Oryza sativa (cultivar:Fujihikari) early- to mid-milking stage
            endosperm tissue_lib:endosperm mRNA, clone_lib:lambda gt11.
  ORGANISM  Oryza sativa
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 2288)
  AUTHORS   Fujita,N., Kubo,A., Francisco,P.B. Jr., Nakakita,M., Harada,K.,
            Minaka,N. and Nakamura,Y.
  TITLE     Purification, characterization, and cDNA structure of isoamylase
            from developing endosperm of rice
  JOURNAL   Planta 208 (2), 283-293 (1999)
  MEDLINE   99266133
REFERENCE   2  (bases 1 to 2288)
  AUTHORS   Nakamura,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUN-1998) Yasunori Nakamura, National Institute of
            Agrobiological Resources, Department of Plant Physiology; 2-1-2
            Kannondai, Tsukuba, Ibaraki 305-8602, Japan
            (E-mail:nakayn@abr.affrc.go.jp, Tel:81-298-38-8382,
            Fax:81-298-38-8347)
FEATURES             Location/Qualifiers
     source          1. .2288
                     /organism="Oryza sativa"
                     /cultivar="Fujihikari"
                     /db_xref="taxon:4530"
                     /chromosome="8"
                     /map="determined"
                     /tissue_type="endosperm"
                     /clone_lib="lambda gt11"
                     /dev_stage="early- to mid-milking stage"
                     /tissue_lib="endosperm"
     CDS             <1. .2202
                     /codon_start=1
                     /product="isoamylase"
                     /protein_id="BAA29041.1"
                     /db_xref="GI:3252794"
                     /translation="VEAVVMPERYALGGACRVLAGMPAPLGATALDGGVNFAVYSAGA
                     SAASLCLFTPDDLEADEVTEEVPLDPLFNRTGNVWHVFIEGELHNMLYGYRFDGMFAP
                     HCGQYFDVSNVVVDPYAKAVISRGEYGVPGPGGDCWPQMAGMIPLPYSTFDWQGDLPL
                     RYPQKDLVIYEMHLRGFTKHSSSNVEHPGTYIGAISKLDYLKELGVNCVELMPCHEFN
                     ELEYFSCSSKMNFWGYSTINFFSPMIRYSSGGIRNCGRDAINEFKTFVREAHKRGIEV
                     IMDVVFNHTAEGNEKGPILSFRGIDNSTYYMLAPKGEFYNYSGCGNTFNCNHPVVREF
                     IVDCLRYWVTEMHVDGFRFDLASIMTRGCSLWDPVNVYGSPVEGDMTTTGTPLATPPL
                     IDMISNDPILGDVKLIAEAWDAGGLYQVGQFPHWKIWSEWNGKYRDIVRQFIKGTDGF
                     AGGFAECLCGSPHLYQAGGRKPWHSINFVCAHDGFTLADLVTYNKKYNSSNGEDNRDG
                     ENHNLSWNCGEEGEFAGLSVKRLRKRQMRNFFVSLMVSQGVPMFYMGDEYGHTKGGNN
                     NTYCHDHYVNYFRWDKKEESSDLQRFCSLMTKFRKQCESLGLADFPTAQRLHWHGHQP
                     GKPDWSETSRFVAFSTKDETKGEIYVAFNASHLPAVVGLPERPGYRWEPLVDTGKPAP
                     YDFLTDDLPDRVTPSTCSLISSTPISTPCSATPPSSLNCSLMIERTKYSIV"
BASE COUNT      587 a    508 c    593 g    600 t
ORIGIN


  Query Match          36.7%;  Score 1101.2;  DB 8;  Length 2288;
  Best Local Similarity   68.1%;  Pred. No. 1.5e-196;
  Matches 1923;  Conservative    0;  Mismatches  273;  Indels  628;  Gaps    3;

Qy     120 gaggacgagccggtggcggaggacaggtacgcgctcggcggcgcgtgcagggtgctcgcc 179
           | ||| | |  ||||  |  ||| ||||||||||| || ||||||||||||||||||||||
Db       1 GTGGAGGCGGTGGTGATGCCGGAGAGGTACGCGCTGGGTGGCGCGTGCAGGGTGCTCGCC 60

Qy     180 ggaatgcccgcgccgctgggcgccaccgcgctcgccggcggggtcaatttcgccgtctat 239
           |||||||||||||||| || |||||||||||| ||||||||||||||||||||||||||||
Db      61 GGAATGCCCGCGCCGCTCGGGGCCACCGCGCTCGACGGCGGGGTCAATTTCGCCGTCTAC 120

Qy     240 tccggcggagccaccgccgcggcgctctgcctcttcacgccagaagatctcaaggcggtg 299
           |||| ||| ||  |||||||| ||||||||||||||| || || |||||| ||||
Db     121 TCCGCCGGCGCATCCGCCGCGTCGCTCTGCCTCTTCACCCCCGACGATCTCGAGGC---- 176

Qy     300 gggttgcctcccgagtagagttcatcagctttgcgtgcgccgcgcgccccttttttgggc 359

Db     177 ------------------------------------------------------------ 176

Qy     360 ctgcaatttaagttttgtactggggcaaatgctgcaggatagggtgaccgaggaggttcc 419
                                                 ||||  |||||| ||||||||||
Db     177 ----------------------------------GGATGAGGTGACTGAGGAGGTTCC 200

Qy     420 ccttgacccccctgatgaatcggaccgggaacgtgtggcatgtcttcatcgaaggcgagct 479
             ||||| || ||| | |||||||| ||||| |||||||| ||||||||||||||||||||
Db     201 GCTTGATCCTCTGTTCAATCGGACGGGGAATGTGTGGCACGTCTTCATCGAAGGCGAGCT 260

Qy     480 gcacaacatgctttacgggtacaggttcgacggcacctttgctcctcactgcgggcacta 539
           |||||||||||| |||||||||||||||||| || |  || || |||||||||| || ||
Db     261 GCACAACATGCTGTACGGGTACAGGTTCGATGGTATGTTCGCCCCTCACTGCGGCCAGTA 320

Qy     540 ccttgatgtttccaatgtcgtggtggatccttatgctaaggcagtgataagccgagggga 599
           | | ||||| |||||||||||||||||||||||||||| ||||||||||||||||||| ||
Db     321 CTTCGATGTCTCCAATGTCGTGGTGGATCCTTATGCCAAGGCAGTGATAAGCCGAGGAGA 380

Qy     600 gtatggtgttccagcgcgtggtaacaattgctggcctcagatggctggcatgatccctct 659
           ||||||||| || |  | ||||  | |||||||||||| ||||||||||||||||||||||
Db     381 GTATGGTGTCCCCGGTCCTGGTGGCGATTGCTGGCCTCAAATGGCTGGCATGATCCCTCT 440

Qy     660 tccatatagcacgtttgattgggaaggcgacctacctctaagatatcctcaaaaggacct 719
           ||| || || |||||||||||| |||| |||||||||| |||||||||| ||||| ||
Db     441 TCCGTACAGTACGTTTGATTGGCAAGGTGACCTACCTCTGAGATATCCTCAGAAGGATCT 500

Qy     720 ggtaatatatgagatgcacttgcgtggattcacgaagcatgattcaagcaatgtagaaca 779
           |||||  |||||||||||  || ||||| || || |||||   ||||||||||||||||||
Db     501 TGTAATCTATGAGATGCATTTACGTGGGTTTACAAAGCACAGTTCAAGCAATGTAGAACA 560

Qy     780 tccgggtactttcattggagctgtgtcgaagcttgactatttgaaggagcttggagttaa 839
           ||| || |||| |||||| ||| | || ||||||||||| ||||||||||||||||||||
Db     561 TCCAGGGACTTACATTGGGGCTATATCAAAGCTTGACTATCTGAAGGAGCTTGGAGTTAA 620
```


Qy 840 ttgtattgaattaatgccctgccatgagttcaacgagctggagtactcaacctcttcttc 899
Db 621 CTGTGTAGAGTTGATGCCCTGCCATGAATTCAATGAGCTGGAGTACTTCAGCTGCTCTTC 680
Qy 900 caagatgaacttttggggatattctaccataaacttcttttcaccaatgacaagatacac 959
Db 681 CAAGATGAACTTCTGGGGATACTCCACGATAAACTTTTTTTCACCAATGATAAGATATTC 740
Qy 960 atcaggcgggataaaaaactgtgggcgtgatgccataaatgagttcaaaacttttgtaag 1019
Db 741 ATCAGGTGGGATAAGAAACTGTGGCCGTGATGCCATAAATGAATTCAAAACTTTTGTTAG 800
Qy 1020 agaggctcacaaacggggaattgaggtgatcctggatgttgtcttcaaccatacagctga 1079
Db 801 AGAGGCTCACAAACGGGGAATTGAGGTGATCATGGATGTTGTCTTCAATCATACAGCCGA 860
Qy 1080 gggtaatgagaatggtccaatattatcatttaaggggtcgataatactacatactatat 1139
Db 861 GGGTAATGAGAAAGGACCAATATTATCATTTAGGGGGATAGATAATAGCACATACTATAT 920
Qy 1140 gcttgcacccaagggagagttttataactattctggctgtgggaataccttcaactgtaa 1199
Db 921 GCTTGCCCCTAAGGGAGAGTTTTACAATTATTCTGGTTGTGGGAATACCTTCAACTGTAA 980
Qy 1200 tcatcctgtggttcgtcaattcattgtagattgtttaagatactgggtgacggaaatgca 1259
Db 981 TCATCCTGTGGTCCGTGAATTTATTGTAGATTGTTTAAGATACTGGGTGACAGAAATGCA 1040
Qy 1260 tgttgatggttttcgttttgatcttgcatccataatgaccagaggttccagtctgtggga 1319
Db 1041 TGTTGATGGTTTTCGTTTTGATCTTGCATCCATAATGACCAGAGGATGCAGTCTTTGGGA 1100
Qy 1320 tccagttaacgtgtatggagctccaatagaaggtgacatgatcacaacagggacacctct 1379
Db 1101 TCCAGTTAATGTGTATGGAAGTCCAGTAGAAGGTGACATGACTACGACAGGGACACCTCT 1160
Qy 1380 tgttactccaccacttattgacatgatcagcaatgacccaattcttggaggcgtcaagct 1439
Db 1161 TGCTACTCCACCACTTATTGACATGATCAGCAATGATCCAATTCTTGGAGATGTCAAGCT 1220
Qy 1440 cattgctgaagcatgggatgcaggaggcctctatcaagtaggtcaattccctcactggaa 1499
Db 1221 CATTGCTGAAGCATGGGATGCGGGAGGCCTCTATCAAGTAGGTCAATTTCCTCACTGGAA 1280
Qy 1500 tgtttggtctgagtggaatgggaagtaccgggacattgtgcgtcaattcattaaaggcac 1559
Db 1281 AATTTGGTCAGAATGGAATGGAAAGTACCGGGACATTGTTCGTCAATTCATTAAAGGCAC 1340
Qy 1560 tgatggatttgctggtggttttgccgaatgtctttgtggaagtccacacctataccaggt 1619
Db 1341 AGATGGATTTGCTGGTGGTTTCGCTGAATGTTTGTGTGGAAGTCCACATCTATAC----- 1395
Qy 1620 aagttgtggcaatacttgtaaatgagttgagtgaatgtcacctggatttttatatatac 1679
Db 1396 ---------- 1395
Qy 1680 cacatgatgatacacatctaaatatataacaatcatagtgtatgcatatgcatttggcta 1739
Db 1396 ---------- 1395
Qy 1740 agaagtattagtgtatacactagtgctatatataggttttaacacccaacttgccaatga 1799
Db 1396 ---------- 1395
Qy 1800 aggaacatagggctttctagttatcttatttatttgtccggtgaataatccactgaaaaa 1859
Db 1396 ---------- 1395
Qy 1860 ttccagccatgtcatttttttaggggggagaagaaactatattgatttgccccccctaaaa 1919
Db 1396 ---------- 1395
Qy 1920 gaagccatctcagaattcataggtaagttgcttttctgtaaagaaaggaaaacgacttca 1979
Db 1396 ---------- 1395
Qy 1980 tactttctatcggtgctaacttagctcgatgtatatttgtaagatgaatgccaaatttaa 2039
Db 1396 ---------- 1395
Qy 2040 tttgtcggataatttgatctgttattcacaaatttctatttggtttctctagaaatcaaa 2099
Db 1396 ---------- 1395
Qy 2100 ccagtaacttgttattggcactgcaacttcttattgattaatcaggcaggaggaaggaaa 2159
Db 1396 ----------CAGGCAGGGGGGAGGAAA 1413
Qy 2160 ccttggcacagtatcaactttgtatgtgcacatgatggatttacactggctgatttggta 2219
Db 1414 CCTTGGCACAGTATCAACTTTGTGTGTGCGCATGATGGATTTACGCTGGCTGATTTGGTG 1473
Qy 2220 acatataataagaagtacaatttaccaaatggggagaacaacagagatggagaaaatcac 2279
Db 1474 ACATATAACAAGAAGTACAACTCTTCAAATGGTGAGGATAATAGAGATGGAGAAAATCAT 1533
Qy 2280 aatcttagctggaattgtggggaggaaggagaattcgcaagattgtctgtcaaaagattg 2339
Db 1534 AACCTCAGCTGGAACTGTGGGGAGGAAGGAGAATTTGCAGGTTTGTCGGTCAAAAGATTG 1593
Qy 2340 aggaagaggcagatgcgcaatttctttgtttgtctcatggtttctcaaggagttccaatg 2399
Db 1594 AGGAAGAGGCAAATGCGCAATTTCTTTGTTTCTCTCATGGTTTCTCAAGGTGTTCCGATG 1653
Qy 2400 ttctacatgggtgatgaatatggccacacaaaaggggcaacaacaacaatacatactgccat 2459
Db 1654 TTCTACATGGGCGATGAATATGGCCATACAAAAGGAGGCAACAACAACACATACTGCCAT 1713
Qy 2460 gattcttatgtcaattatttttcgctgggataaaaaagaacaatactctgagttgcaccga 2519
Db 1714 GACCATTATGTCAATTATTTCCGCTGGGACAAGAAAGAAGAATCCTCTGACTTGCAACGG 1773
Qy 2520 ttctgctgcctcatgaccaaattccgcaaggagtgcgagggtcttggccttgaggacttt 2579
Db 1774 TTCTGCTCTCTTATGACCAAATTCCGCAAACAATGCGAGTCCCTTGGCCTCGCAGACTTC 1833
Qy 2580 ccaacggccaaacggctgcagtggcatggtcatcagcctgggaagcctgattggtctgag 2639
Db 1834 CCAACAGCTCAACGGTTGCATTGGCATGGCCATCAGCCTGGGAAACCTGACTGGTCTGAG 1893
Qy 2640 aatagccgattcgttgccttttccatgaaagatgaaagacagggcgagatctatgtggcc 2699
Db 1894 ACAAGCCGTTTCGTGGCCTTCTCCACGAAAGATGAAACGAAAGGCGAGATCTACGTGGCC 1953
Qy 2700 ttcaacaccagccacttaccggccgttgttgagctcccagagcgcgcagggcgccggtgg 2759
Db 1954 TTCAACGCCAGCCATTTGCCGGCGGTTGTTGGACTCCCAGAGCGCCCAGGCTACCGCTGG 2013
Qy 2760 gaaccggtggtggacacaggcaagccagcaccatacgacttcctcaccgacgacttacct 2819
Db 2014 GAGCCGCTGGTGGACACAGGCAAGCCAGCGCCTTATGACTTCCTCACCGACGACCTGCCT 2073
Qy 2820 gatcgcgctctcaccatacaccagttctcgcatttcctctactccaacctctaccccatg 2879
Db 2074 GATCGTG-TCACGCCGTCCACCTGTTCTCTCATTTCCTCAACTCCAATCTCTACCCCATG 2132
Qy 2880 ctcagctactcatcggtcatcctagtattgcgccctgatgtttgagagaccaatatatac 2939
Db 2133 CTCAGCTACTCCTCCATCATCCTTGAATTGCAGCCTGATGATTGAAAGAACAAAGTATAG 2192
Qy 2940 agta 2943
Db 2193 CATA 2196

RESULT 11

```
AX031277
LOCUS       AX031277               1706 bp    DNA     linear   PAT 20-SEP-2000
DEFINITION  Sequence 16 from Patent WO9914314.
ACCESSION   AX031277
VERSION     AX031277.1  GI:10278607
KEYWORDS    .
SOURCE      Aegilops tauschii.
  ORGANISM  Aegilops tauschii
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Aegilops.
REFERENCE   1  (bases 1 to 1706)
  AUTHORS   Li,Z., Morell,M. and Rahman,S.
  TITLE     Regulation of gene expression in plants
  JOURNAL   Patent: WO 9914314-A 16 25-MAR-1999;
            GOODMAN FIELDER LTD (AU) ; LI ZHONGYI (AU) ; MORELL MATTHEW (AU) ;
            RAHMAN SADEQUR (AU) ; UNIV AUSTRALIAN (AU) ; COMMW SCIENT IND RES
            ORG (AU) ; GROUPE LIMAGRAIN PACIFIC PTY L (AU)
FEATURES             Location/Qualifiers
     source          1. .1706
                     /organism="Aegilops tauschii"
                     /db_xref="taxon:37682"
                     /tissue_type="ENDOSPERM"
     CDS             <1. .1599
                     /codon_start=1
                     /product="PARTIAL CDNA FOR HEXAPLOID WHEAT DBE"
                     /protein_id="CAC09652.1"
                     /db_xref="GI:10278608"
                     /translation="AVSKLDYLKELGVNCIELMPCHEFNELEYSTSSSKMNFWGYSTI
                     NFFSPMTRYTSGGIKNCGRDAINEFKTFVREAHKRGIEVILDVVFNHTAEGNENGPIL
                     SFRGVDNTTYYMLAPKGEFYNYSGCGNTFNCNHPVVRQFIVDCLRYWVMEMHVDGFRF
                     DLASIMTRGSSLWDPVNVYGAPIEGDMITTGTPLVTPPLIDMISNDPILGGVKLIAEA
                     WDAGGLYQVGQFPHWNVWSEWNGKYRDIVRQFIKGTDGFAGGFAECLCGSPHLYQAGG
                     RKPWHSINFVCAHDGFTLGDLVTYNNKYNLPNGENNRDGENHNLSWNCGEEGEFARLS
                     VKRLRKRQMRNFFVCLMVSQGVPMFYMGDEYGHTKGGNNNTYCHDSYVNYFRWDKKEQ
                     YSDLHRFCCLMTKFRKECEGLGLEDFPTAERLQWHGHQPGKPDWSENSRFVAFSMKDE
                     RQGEIYVAFNTSHLPAVVELPERAGRRWEPVVDTGKPAPYDFLTDDLPDRALTIHQFS
                     HFLNSNLYPMLSYSSVILVLRPDV"
BASE COUNT      467 a    355 c    415 g    469 t
ORIGIN
```

```
  Query Match           35.8%;  Score 1074.4;  DB 6;  Length 1706;
  Best Local Similarity  74.9%;  Pred. No. 1.5e-191;
  Matches 1621;  Conservative    0;  Mismatches   16;  Indels  527;  Gaps    1;

Qy    799 gctgtgtcgaagcttgactatttgaaggagcttggagttaattgtattgaattaatgccc 858
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GCTGTGTCGAAGCTTGACTATTTGAAGGAGCTTGGAGTTAATTGTATTGAATTAATGCCC 60

Qy    859 tgccatgagttcaacgagctggagtactcaacctcttcttccaagatgaacttttgggga 918
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TGCCATGAGTTCAACGAGCTGGAGTACTCAACCTCTTCTTCCAAGATGAACTTTTGGGGA 120

Qy    919 tattctaccataaacttcttttcaccaatgacaagatacacatcaggcgggataaaaaaac 978
          |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db    121 TATTCTACCATAAACTTCTTTTCACCAATGACGAGATACACATCAGGCGGGATAAAAAAC 180

Qy    979 tgtgggcgtgatgccataaatgagttcaaaacttttgtaagagaggctcacaaacgcggga 1038
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TGTGGGCGTGATGCCATAAATGAGTTCAAAACTTTTGTAAGAGAGGCTCACAAACGCGGGA 240

Qy   1039 attgaggtgatcctggatgttgtcttcaaccatacagctgagggtaatgagaatggtccca 1098
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 ATTGAGGTGATCCTGGATGTTGTCTTCAACCATACAGCTGAGGGTAATGAGAATGGTCCA 300

Qy   1099 atattatcatttaaggggtcgataatactacatactatatgcttgcacccaagggagag 1158
          |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db    301 ATATTATCATTTAGGGCGGTCGATAATACTACATACTATATGCTTGCACCCAAGGGAGAG 360

Qy   1159 ttttataactattctggctgtgggaatacctt caactgtaatcatcctgtggttcgtcaa 1218
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TTTTATAACTATTCTGGCTGTGGGAATACCTTCAACTGTAATCATCCTGTGGTTCGTCAA 420

Qy   1219 ttcattgtagattgtttaagatactgggtgacggaaatgcatgttgatggttttcgtttt 1278
          ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db    421 TTCATTGTAGATTGTTTAAGATACTGGGTGATGGAAATGCATGTTGATGGTTTTCGTTTT 480

Qy   1279 gatcttgcatccataatgaccagaggttccagtctgtgggatccagttaacgtgtatgga 1338
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 GATCTTGCATCCATAATGACCAGAGGTTCCAGTCTGTGGGATCCAGTTAACGTGTATGGA 540

Qy   1339 gctccaatagaaggtgacatgatcacaacagggacacctcttgttactccaccacttatt 1398
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GCTCCAATAGAAGGTGACATGATCACAACAGGGACACCTCTTGTTACTCCACCACTTATT 600

Qy   1399 gacatgatcagcaatgacccaattcttggaggcgtcaagctcattgctgaagcatgggat 1458
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 GACATGATCAGCAATGACCCAATTCTTGGAGGCGTCAAGCTCATTGCTGAAGCATGGGAT 660

Qy   1459 gcaggaggcctctatcaagtaggtcaattccctcactggaatgtttggtctgagtggaat 1518
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 GCAGGAGGCCTCTATCAAGTAGGTCAATTCCCTCACTGGAATGTTTGGTCTGAGTGGAAT 720

Qy   1519 gggaagtaccgggacattgtgcgtcaattcattaaaggcactgatggatttgctggtggt 1578
          |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db    721 GGGAAGTACCGGGACATTGTGCGCCAATTCATTAAAGGCACTGATGGATTTGCTGGTGGT 780

Qy   1579 tttgccgaatgtctttgtggaagtccacacctataccaggtaagttgtggcaatacttgt 1638
          ||||||||||||||||||||||||||||||||||
Db    781 TTTGCCGAATGTCTTTGTGGAAGTCCACACCTATAC---------------------- 816

Qy   1639 aaatgagttgagtgaatgtcacctggatttttatatataccacatgatgatacacatct 1698
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    817 ---------------------------------------------------------- 816

Qy   1699 aaatatataacaatcatagtgtatgcatatgcatttggctaagaagtattagtgtataca 1758
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    817 ---------------------------------------------------------- 816

Qy   1759 ctagtgctatatataggttttaacacccaacttgccaatgaaggaacatagggctttcta 1818
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    817 ---------------------------------------------------------- 816

Qy   1819 gttatcttatttatttgtccggtgaataatccactgaaaaattccagccatgtcatttttt 1878
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    817 ---------------------------------------------------------- 816

Qy   1879 tagggggggagaagaaactatattgatttgccccccctaaaagaagccatctcagaattca 1938
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    817 ---------------------------------------------------------- 816

Qy   1939 taggtaagttgcttttctgtaaagaaaggaaaacgacttcatactttctatcggtgctaa 1998
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    817 ---------------------------------------------------------- 816

Qy   1999 cttagctcgatgtatatttgtaagatgaatgccaaatttaatttgtcggataatttgatc 2058
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    817 ---------------------------------------------------------- 816

Qy   2059 tgttattcacaaatttctatttggtttctctagaaatcaaaccagtaacttgttattggc 2118
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    817 ---------------------------------------------------------- 816

Qy   2119 actgcaacttcttattgattaatcaggcaggaggaaggaaaccttggcacagtatcaact 2178
                                  ||||||||||||||||||||||||||||||||||||
Db    817 ----------------------CAGGCAGGAGGAAGGAAACCTTGGCACAGTATCAACT 853

Qy   2179 ttgtatgtgcacatgatggatttacactggctgatttggtaacatataataagaagtaca 2238
          |||||||||||||||||||||||||||||||||| |||||||||||||||||| |||||||
Db    854 TTGTATGTGCACATGATGGATTTACACTGGGTGATTTGGTAACATATAATAACAAGTACA 913

Qy   2239 atttaccaaatggggagaacaacagagatggagaaaatcacaatcttagctggaattgtg 2298
          |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db    914 ATTTACCAAATGGGGAGAACAATAGAGATGGAGAAAATCACAATCTTAGCTGGAATTGTG 973
```

```
Qy   2299 gggaggaaggagaattcgcaagattgtctgtcaaaagattgaggaagaggcagatgcgca 2358
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    974 GGGAGGAAGGAGAATTCGCAAGATTGTCTGTCAAAAGATTGAGGAAGAGGCAGATGCGCA 1033

Qy   2359 atttctttgtttgtctcatggtttctcaaggagttccaatgttctacatgggtgatgaat 2418
          ||||||||||||||||||||||||||||||||||||||||| |||||||| |||||||
Db   1034 ATTTCTTTGTTTGTCTCATGGTTTCTCAAGGAGTTCCAATGTTTTACATGGGCGATGAAT 1093

Qy   2419 atggccacacaaaaggggcaacaacaatacatactgccatgattcttatgtcaattatt 2478
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1094 ATGGCCACACAAAAGGGGCAACAACAATACATACTGCCATGATTCTTATGTCAATTATT 1153

Qy   2479 ttcgctgggataaaaaagaacaatactctgagttgcaccgattctgctgcctcatgacca 2538
          ||||||||||||||||||||||||||||||| |||||| ||||||||||||||||||||
Db   1154 TTCGCTGGGATAAAAAAGAACAATACTCTGACTTGCACAGATTCTGCTGCCTCATGACCA 1213

Qy   2539 aattccgcaaggagtgcgagggtcttggccttgaggactttccaacggccaaacggctgc 2598
          |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db   1214 AATTCCGCAAGGAGTGCGAGGGTCTTGGCCTTGAGGACTTTCCAACGGCCGAACGGCTGC 1273

Qy   2599 agtggcatggtcatcagcctgggaagcctgattggtctgagaatagccgattcgttgcct 2658
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1274 AGTGGCATGGTCATCAGCCTGGGAAGCCTGATTGGTCTGAGAATAGCCGATTCGTTGCCT 1333

Qy   2659 tttccatgaaagatgaaagacagggcgagatctatgtggccttcaacaccagccacttac 2718
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1334 TTTCCATGAAAGATGAAAGACAGGGCGAGATCTATGTGGCCTTCAACACCAGCCACTTAC 1393

Qy   2719 cggccgttgttgagctcccagagcgcgcagggcgccggtgggaaccggtggtggacacag 2778
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1394 CGGCCGTTGTTGAGCTCCCAGAGCGCGCAGGGCGCCGGTGGGAACCGGTGGTGGACACAG 1453

Qy   2779 gcaagccagcaccatacgacttcctcaccgacgacttacctgatcgcgctctcaccatac 2838
          |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db   1454 GCAAGCCAGCACCATATGACTTCCTCACCGACGACTTACCTGATCGCGCTCTCACCATAC 1513

Qy   2839 accagttctcgcatttcctctactccaacctctaccccatgctcagctactcatcggtca 2898
          |||||||||| |||||||| ||||||||||||||||||||||||||||||||||||||||
Db   1514 ACCAGTTCTCTCATTTCCTCAACTCCAACCTCTACCCCATGCTCAGCTACTCATCGGTCA 1573

Qy   2899 tcctagtattgcgccctgatgtttgagagaccaatatatacagtaaataatatgtctata 2958
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db   1574 TCCTAGTATTGCGCCCTGATGTTTGAGAGACAAATATATACAGTAAATAATATGTCTATA 1633

Qy   2959 tgta 2962
          ||||
Db   1634 TGTA 1637
```

```
RESULT  12
AR072143
LOCUS       AR072143               2712 bp    DNA     linear   PAT 18-FEB-2000
DEFINITION  Sequence 1 from patent US 5912413.
ACCESSION   AR072143
VERSION     AR072143.1  GI:7223031
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 2712)
  AUTHORS   Myers,A.M. and James,M.Graham.
  TITLE     Isolation of SU1, a starch debranching enzyme, the product of the
            maize gene sugary1
  JOURNAL   Patent: US 5912413-A 1 15-JUN-1999;
FEATURES             Location/Qualifiers
     source          1. .2712
                     /organism="unknown"
BASE COUNT      636 a    646 c    745 g    685 t
ORIGIN
```

```
  Query Match           33.4%;  Score 1000;  DB 6;  Length 2712;
  Best Local Similarity 65.7%;  Pred. No. 1.5e-177;
```

```
  Matches 1883;  Conservative    0;  Mismatches  350;  Indels  633;  Gaps    4;

Qy     79 cggttgtcgaggcggcgacgaaggtagaggacgaggggggaggaggacgagccggtggcgg 138
          | || | |||||||| |  |  ||  |||||||| |  || || ||||||  |||||| |
Db    233 CCGTGGCCGAGGCGGTGCAGGCGGAGGAGGACGACGACGACGACGACGAGGAGGTGGCCG 292

Qy    139 aggacaggtacgcgctcggcggcgcgtgcagggtgctcgccggaatgcccgcgccgctgg 198
          |||| |||| |||||| |||||||||||| |||||||||| ||||||||||||||||| |
Db    293 AGGAGAGGTTCGCGCTGGGCGGCGCGTGCCGGGTGCTCGCGGGAATGCCCGCGCCGCTCG 352

Qy    199 gcgccaccgcgctcgccggcggggtcaatttcgccgtctattccggcggagccaccgccg 258
          |||||||||||||  |||||| ||||| |||||||||| ||| |||| ||| ||||||
Db    353 GCGCCACCGCGCTCCGCGGCGGTGTCAACTTCGCCGTCTACTCCAGCGGTGCCTCCGCCG 412

Qy    259 cggcgctctgcctcttcacgccagaagatctcaaggcggtggggttgcctcccgagtaga 318
          || |||| |||||||| | || |  || ||||||||
Db    413 CGTCGCTGAGCCTCTTCGCTCCCGGCGACCTCAAGGC---------------------- 449

Qy    319 gttcatcagctttgcgtgcgccgcgcgccccttttttgggcctgcaatttaagttttgta 378

Db    450 ---------------------------------------------------------- 449

Qy    379 ctggggcaaatgctgcaggatagggtgaccgaggaggttccccttgaccccctgatgaat 438
                               |||||||||||||||||||| ||||| || |||||| | ||
Db    450 -----------------GGATAGGGTGACCGAGGAGGTGCCCCTCGATCCCCTGCTCAAC 492

Qy    439 cggaccgggaacgtgtggcatgtcttcatcgaagg---cgagctgcacaacatgctttac 495
          || || || |||||||||| || |||||| | ||   |||||||||  |||||| | |
Db    493 CGAACGGGAAACGTGTGGCACGTGTTCATCCACGGGGACGAGCTGCACGGCATGCTCTGC 552

Qy    496 gggtacaggttcgacggcacctttgctcctcactgcgggcactaccttgatgtttccaat 555
          || |||||||||| |||   || || ||| |  |||| || |||   ||||| |||||
Db    553 GGATACAGGTTCGATGGCGTGTTCGCCCCTGAGCGCGGACAGTACTACGATGTGTCCAAC 612

Qy    556 gtcgtggtggatccttatgctaaggcagtgataagccgaggggagtatggtgttccagcg 615
          || |||||||||| || |||||||||||| |||||||||| || |||||||| || |||
Db    613 GTTGTGGTGGATCCATACGCTAAGGCAGTGGTAAGCCGAGGTGAATATGGTGTGCCTGCG 672

Qy    616 cgtggtaacaattgctggcctcagatggctggcatgatccctcttccatatagcacgttt 675
          | ||||   | |||  |||||||| |||||||| |||||||||||||| ||||  | ||||
Db    673 CCTGGTGGTAGTTGTTGGCCTCAAATGGCTGGTATGATCCCTCTTCCCTATAATAAGTTT 732

Qy    676 gattgggaaggcgacctacctctaagatatcctcaaaaggacctggtaatatatgagatg 735
          |||||| |||| |||||||| ||  | || | ||| |||||||| || |||||||| |||
Db    733 GATTGGCAAGGTGACCTACCCCTTGGGTACCATCAGAAGGACCTTGTCATATATGAAATG 792

Qy    736 cacttgcgtggattcacgaagcatgattcaagcaatgtagaacatccgggtactttcatt 795
          || |||||||||||||| |||||  | ||||||||   | |||| || || |||| ||||
Db    793 CATTTGCGTGGATTCACAAAGCACAACTCAAGCAAGACAAAACACCCAGGAACTTACATT 852

Qy    796 ggagctgtgtcgaagcttgactatttgaaggagcttggagttaattgtattgaattaatg 855
          || |||||||| |||||||| || | ||||| |||||||| || ||||| ||  |||||
Db    853 GGTGCTGTGTCAAAGCTTGACCATCTAAAGGAACTTGGAGTGAACTGTATAGAGCTAATG 912

Qy    856 ccctgccatgagttcaacgagctggagtactcaacctcttcttccaagatgaacttttgg 915
          |||||||||||||||||| ||||| |||||||  | ||| |||||  |||||||||| |||
Db    913 CCCTGCCATGAGTTCAATGAGCTAGAGTACTTCAGCTCCTCTTCGAAGATGAACTTCTGG 972

Qy    916 ggatattctaccataaacttcttttcaccaatgacaagatacacatcaggcgggataaaa 975
          |||||||| || ||||| || || ||||||||| |||||||  | ||| | || |||| |
Db    973 GGATATTCCACAATAAATTTTTTCTCACCAATGGCAAGATATTCTTCAAGTGGCATAAGA 1032

Qy    976 aactgtgggcgtgatgccataaatgagttcaaaacttttgtaagagaggctcacaaacgg 1035
           ||| |||  ||| |||||||||||| || ||| |||||||||| ||||| |||||||||
Db   1033 GACTCTGGATGTGGTGCCATAAATGAATTTAAAGCTTTTGTAAGGGAGGCCCACAAACGG 1092

Qy   1036 ggaattgaggtgatcctggatgttgtcttcaaccatacagctgagggtaatgagaatggt 1095
          |||||||||||||||| |||||||||||||||| |||||||||| ||||||||||  ||
Db   1093 GGAATTGAGGTGATCATGGATGTTGTCTTCAATCATACAGCTGAAGGTAATGAGAAAGGC 1152

Qy   1096 ccaatattatcatttaaggggtcgataatactacatactatatgcttgcacccaaggga 1155
          |||||||||| |||| |||| | ||||||| ||||||||| |||||||||| ||||||
Db   1153 CCAATATTATCCTTTAGGGGGATAGATAATAGTACATACTACATGCTTGCACCTAAGGGA 1212
```

```
Qy   1156 gagttttataactattctggctgtgggaataccttcaactgtaatcatcctgtggttcgt 1215
          |||||||||||| |||||||| ||||| |||||||||| |||||||||||||| || |||
Db   1213 GAGTTTTATAATTATTCTGGTTGTGGAAATACCTTCAATTGTAATCATCCTGTAGTCCGT 1272

Qy   1216 caattcattgtagattgtttaagatactgggtgacggaaatgcatgttgatggttttcgt 1275
          |||| || || ||||| || |||||||||||| || ||||||||||||||||||||||||
Db   1273 GAATTTATAGTGGATTGCTTGAGATACTGGGTAACAGAAATGCATGTTGATGGTTTTCGT 1332

Qy   1276 tttgatcttgcatccataatgaccagaggttccagtctgtgggatccagttaacgtgtat 1335
          ||||| |||||||| ||| |||||||||| | |||||| |||||||||||||| ||||||
Db   1333 TTTGACCTTGCATCTATACTGACCAGAGGATGCAGTCTATGGGATCCAGTTAATGTGTAT 1392

Qy   1336 ggagctccaatagaaggtgacatgatcacaacagggacacctcttgttactccaccactt 1395
          |||  |||||| |||||||||||||| || |||||||||||||||| | |||||||||
Db   1393 GGAAGTCCAATGGAAGGTGACATGATTACGACAGGGACACCTCTTGTTGCCCCACCACTT 1452

Qy   1396 attgacatgatcagcaatgacccaattcttggaggcgtcaagctcattgctgaagcatgg 1455
          |||||||||| ||||||||||||||||||||   ||||||||||||||||||||||||
Db   1453 ATTGACATGATTAGCAATGACCCAATTCTTGGAAATGTCAAGCTCATTGCTGAAGCATGG 1512

Qy   1456 gatgcaggaggcctctatcaagtaggtcaattccctcactggaatgtttggtctgagtgg 1515
          |||||||||| |||||||||| |||||| || |||||||||| |||||||| ||||||
Db   1513 GATGCAGGAGGTCTCTATCAAGAAGGTCAGTTTCCTCACTGGAACGTTTGGTCAGAGTGG 1572

Qy   1516 aatgggaagtaccgggacattgtgcgtcaattcattaaaggcactgatggatttgctggt 1575
          ||||| ||||| || || |  |||||||| ||||| |||||||| ||||||||||||||||
Db   1573 AATGGAAAGTATCGCGATACCGTGCGTCAGTTCATCAAAGGCACAGATGGATTTGCTGGT 1632

Qy   1576 ggttttgccgaatgtctttgtggaagtccacacctataccaggtaagttgtggcaatact 1635
          | |||||| ||||| || ||||||||||||||   |||||
Db   1633 GCTTTTGCTGAATGCCTATGTGGAAGTCCACAGTTATAC---------------------- 1671

Qy   1636 tgtaaatgagttgagtgaatgtcacctggatttttttatatataccacatgatgatacaca 1695

Db   1672 ------------------------------------------------------------ 1671

Qy   1696 tctaaatatataacaatcatagtgtatgcatatgcatttggctaagaagtattagtgtat 1755

Db   1672 ------------------------------------------------------------ 1671

Qy   1756 acactagtgctatatataggttttaacacccaacttgccaatgaaggaacatagggcttt 1815

Db   1672 ------------------------------------------------------------ 1671

Qy   1816 ctagttatcttatttatttgtccggtgaataatccactgaaaaattccagccatgtcatt 1875

Db   1672 ------------------------------------------------------------ 1671

Qy   1876 ttttaggggggggagaagaaactatattgatttgccccctaaaagaagccatctcagaat 1935

Db   1672 ------------------------------------------------------------ 1671

Qy   1936 tcataggtaagttgctttttctgtaaagaaaggaaaacgacttcatactttctatcggtgc 1995

Db   1672 ------------------------------------------------------------ 1671

Qy   1996 taacttagctcgatgtatatttgtaagatgaatgccaaatttaatttgtcggataatttg 2055

Db   1672 ------------------------------------------------------------ 1671

Qy   2056 atctgttattcacaaatttctatttggtttctctagaaatcaaaccagtaacttgttatt 2115

Db   1672 ------------------------------------------------------------ 1671

Qy   2116 ggcactgcaacttcttattgattaatcaggcaggaggaaggaaaccttggcacagtatca 2175
                                     |||||||| || ||||| ||||||||||||||
Db   1672 -------------------------CAGGCAGGGGGGAGGAAGCCTTGGCACAGTATCG 1705

Qy   2176 actttgtatgtgcacatgatggatttacactggctgatttggtaacatataataagaagt 2235
          |||||||||||||| |||||||||||||||||||||||||||| ||||| ||||  ||||
Db   1706 GCTTTGTATGTGCACACGATGGATTTACACTGGCTGATTTGGTCACATACAATAGCAAGT 1765
```

```
Qy   2236 acaatttaccaaatggggagaacaacagagatggagaaaatcacaatcttagctggaatt 2295
          |||| ||  |||||||| ||| ||  |||||||||| |||||||| ||||||||||||||||
Db   1766 ACAACTTGTCAAATGGTGAGGACTTCAGAGATGGGGAAAATCATAATCTTAGCTGGAATT 1825

Qy   2296 gtggggaggaaggagaattcgcaagattgtctgtcaaaagattgaggaagaggcagatgc 2355
          |||||||||||||||||||| |||||  |||| |||  |||||| |||||||||| ||||
Db   1826 GTGGGGAGGAAGGAGAATTTGCAAGTCTGTCAGTCCGAAGATTAAGGAAGAGGCAAATGC 1885

Qy   2356 gcaatttctttgtttgtctcatggtttctcaaggagttccaatgttctacatgggtgatg 2415
          |||||||||||||||||||| |||||||||| |||||||||||||||||||||| ||||
Db   1886 GCAATTTCTTTGTTTGTCTTATGGTTTCTCAGGGAGTTCCAATGTTCTACATGGGCGATG 1945

Qy   2416 aatatggccacacaaaaggggcaacaacaatacatactgccatgattcttatgtcaatt 2475
          ||||||| |||||||| || || |||||||||| ||||||||||   ||||||||||
Db   1946 AATATGGTCACACAAAGGGAGGGAACAACAATACGTACTGCCATGACCATTATGTCAATT 2005

Qy   2476 attttcgctgggataaaa---aagaacaatactctgagttgcaccgattctgctgcctca 2532
          |||| || ||||||||| |   ||||||||| |||||| ||| || |||||||| | ||||
Db   2006 ATTTCCGTTGGGATAAGAAGGAAGAACAATCCTCTGATTTGTACAGATTCTGCCGTCTCA 2065

Qy   2533 tgaccaaattccgcaaggagtgcgagggtcttggccttgaggactttccaacggccaaac 2592
          ||||| |||||||||| || || ||   |||||||||||||||||| || ||  |  |||
Db   2066 TGACCGAATTCCGCAAAGAATGTGAATCTCTTGGCCTTGAGGACTTCCCGACTTCAGAAC 2125

Qy   2593 ggctgcagtggcatggtcatcagcctgggaagcctgattggtctgagaatagccgattcg 2652
          || || | ||||| |||||||||| |||||||||| ||||| |||   ||||||||||
Db   2126 GGTTGAAATGGCACGGTCATCAGCCCGGGAAGCCTGACTGGTCAGAGGCAAGCCGATTCG 2185

Qy   2653 ttgccttttccatgaaagatgaaagacagggcgagatctatgtggccttcaacaccagcc 2712
          |||||||  ||||||| || ||||   | |||||||||| |||||||||||||||| |
Db   2186 TTGCCTTCACCATGAAGGACGAAACCAAAGGCGAGATCTACGTGGCCTTCAACACCAGTC 2245

Qy   2713 acttaccggccgttgttgagctcccagagcgcgcagggcgccggtgggaaccggtggtgg 2772
          || | ||||  ||||||| ||| |||||||| | |||  ||| ||||| ||||||||||
Db   2246 ACCTTCCGGTGGTTGTTGGGCTTCCAGAGCGCTCTGGGTTCCGATGGGAGCCGGTGGTGG 2305

Qy   2773 acacaggcaagccagcaccatacgacttcctcaccgacgacttacctgatcgcgctctca 2832
          |||| ||||||   |||||||| |||||||||||||| | | | || ||||| ||| |||
Db   2306 ACACCGGCAAGGAGGCACCATATGACTTCCTCACCGATGGCCTGCCAGATCGTGCTGTCA 2365

Qy   2833 ccatacaccagttctcgcatttcctctactccaacctctaccccatgctcagctactcat 2892
          || |  |||||||||| |||||||| ||||||| ||||| || |||||||||||||| |
Db   2366 CCGTCTACCAGTTCTCTCATTTCCTCAACTCCAATCTCTATCCTATGCTCAGCTACTCCT 2425

Qy   2893 cggtcatcctagtattgcgccctgatgtttgagagaccaatatata 2938
          |  ||||||| |||||||||||||||| ||| |||   | ||| |
Db   2426 CCATCATCCTTGTATTGCGCCCTGATGTCTGAAAGAAGCAGATACA 2471
```

RESULT 13
ZMU18908

LOCUS	ZMU18908 2712 bp mRNA linear PLN 06-MAR-1996
DEFINITION	Zea mays Sulp (Sugary1) mRNA, partial cds.
ACCESSION	U18908
VERSION	U18908.1 GI:758725
KEYWORDS	.
SOURCE	maize.
ORGANISM	Zea mays Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.
REFERENCE	1 (bases 1 to 2712)
AUTHORS	James,M.G., Robertson,D.S. and Myers,A.M.
TITLE	Characterization of the maize gene sugary1, a determinant of starch composition in kernels
JOURNAL	Plant Cell 7 (4), 417-429 (1995)
MEDLINE	95290995
REFERENCE	2 (bases 1 to 2712)
AUTHORS	Myers,A.M.
TITLE	Direct Submission
JOURNAL	Submitted (19-DEC-1994) Alan M. Myers, Biochemistry and Biophysics, Iowa State University, 2110 Molecular Biology Building, Ames, IA 50011, USA

```
COMMENT     On Apr 6, 1995 this sequence version replaced gi:755295.
FEATURES             Location/Qualifiers
     source          1. .2712
                     /organism="Zea mays"
                     /db_xref="taxon:4577"
                     /chromosome="4S"
                     /clone="lambda clone H-4; subclones pMJ67 and pMJ68; PCR
                     product clone pMJ125"
                     /tissue_type="endosperm"
                     /clone_lib="lambda gt11 cDNA library from R. Schmidt,
                     Dept. of Biology, University of California at San Diego,
                     La Jolla, CA 92093, USA"
                     /dev_stage="20 days after pollination"
     gene            1. .2712
                     /gene="Sugary1"
     CDS             <1. .2457
                     /gene="Sugary1"
                     /note="similar to Pseudomonas sp. isoamylase, Swiss-Prot
                     Accession Number P26501"
                     /codon_start=1
                     /product="Sulp"
                     /protein_id="AAA91298.1"
                     /db_xref="GI:758726"
                     /translation="RLVTHSTRTHYLIGQSQTNWAPSPPLPLPMAQKLPCVSSPRPLL
                     AVPAGRWRAGVRGRPNVAGLGRGRLSLHAAAARPVAEAVQAEEDDDDDDEEVAEERFA
                     LGGACRVLAGMPAPLGATALRGGVNFAVYSSGASAASLSLFAPGDLKADRVTEEVPLD
                     PLLNRTGNVWHVFIHGDELHGMLCGYRFDGVFAPERGQYYDVSNVVVDPYAKAVVSRG
                     EYGVPAPGGSCWPQMAGMIPLPYNKFDWQGDLPLGYHQKDLVIYEMHLRGFTKHNSSK
                     TKHPGTYIGAVSKLDHLKELGVNCIELMPCHEFNELEYFSSSSKMNFWGYSTINFFSP
                     MARYSSSGIRDSGCGAINEFKAFVREAHKRGIEVIMDVVFNHTAEGNEKGPILSFRGI
                     DNSTYYMLAPKGEFYNYSGCGNTFNCNHPVVREFIVDCLRYWVTEMHVDGFRFDLASI
                     LTRGCSLWDPVNVYGSPMEGDMITTGTPLVAPPLIDMISNDPILGNVKLIAEAWDAGG
                     LYQEGQFPHWNVWSEWNGKYRDTVRQFIKGTDGFAGAFAECLCGSPQLYQAGGRKPWH
                     SIGFVCAHDGFTLADLVTYNSKYNLSNGEDFRDGENHNLSWNCGEEGEFASLSVRRLR
                     KRQMRNFFVCLMVSQGVPMFYMGDEYGHTKGGNNNTYCHDHYVNYFRWDKKEEQSSDL
                     YRFCRLMTEFRKECESLGLEDFPTSERLKWHGHQPGKPDWSEASRFVAFTMKDETKGE
                     IYVAFNTSHLPVVVGLPERSGFRWEPVVDTGKEAPYDFLTDGLPDRAVTVYQFSHFLN
                     SNLYPMLSYSSIILVLRPDV"
BASE COUNT      637 a    644 c    745 g    686 t
ORIGIN
```

```
  Query Match             33.4%;  Score 1000;  DB 8;  Length 2712;
  Best Local Similarity   65.7%;  Pred. No. 1.5e-177;
  Matches 1883;  Conservative    0;  Mismatches  350;  Indels  633;  Gaps    4;

Qy      79 cggttgtcgaggcggcgacgaaggtagaggacgaggggggaggaggacgagccggtggcgg 138
           | || | |||||||| | |  | ||  |||||||| |  || || ||||||  |||||| |
Db     233 CCGTGGCCGAGGCGGTGCAGGCGGAGGAGGACGACGACGACGACGACGAGGAGGTGGCCG 292

Qy     139 aggacaggtacgcgctcggcggcgcgtgcagggtgctcgccggaatgcccgcgccgctgg 198
           |||| |||| |||||| |||||||||||| |||||||||| ||||||||||||||||||| |
Db     293 AGGAGAGGTTCGCGCTGGGCGGCGCGTGCCGGGTGCTCGCGGGAATGCCCGCGCCGCTCG 352

Qy     199 gcgccaccgcgctcgccggcggggtcaatttcgccgtctattccggcggagccaccgccg 258
           ||||||||||||  |||||| |||| |||||||||| ||| |||| ||| ||||||
Db     353 GCGCCACCGCGCTCCGCGGCGGTGTCAACTTCGCCGTCTACTCCAGCGGTGCCTCCGCCG 412

Qy     259 cggcgctctgcctcttcacgccagaagatctcaaggcggtggggttgcctcccgagtaga 318
           || ||||  |||||||| | || |  || ||||||||
Db     413 CGTCGCTGAGCCTCTTCGCTCCCGGCGACCTCAAGGC---------------------- 449

Qy     319 gttcatcagctttgcgtgcgccgcgcgccccttttttgggcctgcaatttaagttttgta 378

Db     450 ------------------------------------------------------------ 449

Qy     379 ctggggcaaatgctgcaggatagggtgaccgaggaggttcccccttgaccccctgatgaat 438
                               |||||||||||||||||||| ||||| || |||||| | ||
Db     450 -----------------GGATAGGGTGACCGAGGAGGTGCCCCTCGATCCCCTGCTCAAC 492

Qy     439 cggaccgggaacgtgtggcatgtcttcatcgaagg---cgagctgcacaacatgctttac 495
           || || || |||||||||||| || |||||| | ||   ||||||||||  |||||| | |
Db     493 CGAACGGGAAACGTGTGGCACGTGTTCATCCACGGGGACGAGCTGCACGGCATGCTCTGC 552
```

```
Qy     496 gggtacaggttcgacggcacctttgctcctcactgcgggcactaccttgatgtttccaat 555
           || |||||||||||| ||| || || ||| | |||| || |||  ||||| |||||
Db     553 GGATACAGGTTCGATGGCGTGTTCGCCCCTGAGCGCGGACAGTACTACGATGTGTCCAAC 612

Qy     556 gtcgtggtggatccttatgctaaggcagtgataagccgaggggagtatggtgttccagcg 615
           || |||||||||| || |||||||||| |||||||| || |||||||| || |||
Db     613 GTTGTGGTGGATCCATACGCTAAGGCAGTGGTAAGCCGAGGTGAATATGGTGTGCCTGCG 672

Qy     616 cgtggtaacaattgctggcctcagatggctggcatgatccctcttccatatagcacgttt 675
           | ||||   | ||| |||||||| |||||||| |||||||||||||| ||||  | ||||
Db     673 CCTGGTGGTAGTTGTTGGCCTCAAATGGCTGGTATGATCCCTCTTCCCTATAATAAGTTT 732

Qy     676 gattgggaaggcgacctacctctaagatatcctcaaaaggacctggtaatatatgagatg 735
           |||||| |||| |||||||| ||  | || | ||| |||||||| || |||||||| |||
Db     733 GATTGGCAAGGTGACCTACCCCTTGGGTACCATCAGAAGGACCTTGTCATATATGAAATG 792

Qy     736 cacttgcgtggattcacgaagcatgattcaagcaatgtagaacatccgggtactttcatt 795
           || |||||||||||||| |||||  | |||||||   | |||| || || |||| ||||
Db     793 CATTTGCGTGGATTCACAAAGCACAACTCAAGCAAGACAAAACACCCAGGAACTTACATT 852

Qy     796 ggagctgtgtcgaagcttgactatttgaaggagcttggagttaattgtattgaattaatg 855
           || |||||||| |||||||| || | |||| |||||||| || |||| ||  ||||
Db     853 GGTGCTGTGTCAAAGCTTGACCATCTAAAGGAACTTGGAGTGAACTGTATAGAGCTAATG 912

Qy     856 ccctgccatgagttcaacgagctggagtactcaacctcttcttccaagatgaacttttgg 915
           |||||||||||||||| |||| |||||||  | ||| |||| |||||||||| |||
Db     913 CCCTGCCATGAGTTCAATGAGCTAGAGTACTTCAGCTCCTCTTCGAAGATGAACTTCTGG 972

Qy     916 ggatattctaccataaacttcttttcaccaatgacaagatacacatcaggcgggataaaa 975
           |||||||| || |||| || || |||||||| |||||||  | ||| | || |||| |
Db     973 GGATATTCCACAATAAATTTTTTCTCACCAATGGCAAGATATTCTTCAAGTGGCATAAGA 1032

Qy     976 aactgtgggcgtgatgccataaatgagttcaaaacttttgtaagagaggctcacaaacgg 1035
           ||| |||  ||| |||||||||||| || ||| |||||||||| |||| |||||||||
Db    1033 GACTCTGGATGTGGTGCCATAAATGAATTTAAAGCTTTTGTAAGGGAGGCCCACAAACGG 1092

Qy    1036 ggaattgaggtgatcctggatgttgtcttcaaccatacagctgagggtaatgagaatggt 1095
           |||||||||||||| |||||||||||||| |||||||||| |||||||||| ||
Db    1093 GGAATTGAGGTGATCATGGATGTTGTCTTCAATCATACAGCTGAAGGTAATGAGAAAGGC 1152

Qy    1096 ccaatattatcatttaaggggtcgataatactacatactatatgcttgcacccaaggga 1155
           |||||||||| |||| |||| | ||||||| |||||||| |||||||||| ||||||
Db    1153 CCAATATTATCCTTTAGGGGGATAGATAATAGTACATACTACATGCTTGCACCTAAGGGA 1212

Qy    1156 gagttttataactattctggctgtgggaatacctt caactgtaatcatcctgtggttcgt 1215
           |||||||||| |||||||| ||||| |||||||||| |||||||||||||| || |||
Db    1213 GAGTTTTATAATTATTCTGGTTGTGGAAATACCTTCAATTGTAATCATCCTGTAGTCCGT 1272

Qy    1216 caattcattgtagattgtttaagatactgggtgacggaaatgcatgttgatggttttcgt 1275
           |||| || || ||||| || |||||||||| || ||||||||||||||||||||||||
Db    1273 GAATTTATAGTGGATTGCTTGAGATACTGGGTAACAGAAATGCATGTTGATGGTTTTCGT 1332

Qy    1276 tttgatcttgcatccataatgaccagaggttccagtctgtgggatccagttaacgtgtat 1335
           ||||| |||||||| ||| |||||||||| | |||||| |||||||||||||| ||||||
Db    1333 TTTGACCTTGCATCTATACTGACCAGAGGATGCAGTCTATGGGATCCAGTTAATGTGTAT 1392

Qy    1336 ggagctccaatagaaggtgacatgatcacaacagggacacctcttgttactccaccactt 1395
           |||  |||||| |||||||||||||| || ||||||||||||||||| | |||||||||
Db    1393 GGAAGTCCAATGGAAGGTGACATGATTACGACAGGGACACCTCTTGTTGCCCCACCACTT 1452

Qy    1396 attgacatgatcagcaatgacccaattcttggaggcgtcaagctcattgctgaagcatgg 1455
           |||||||||| |||||||||||||||||||   ||||||||||||||||||||||||||
Db    1453 ATTGACATGATTAGCAATGACCCAATTCTTGGAAATGTCAAGCTCATTGCTGAAGCATGG 1512

Qy    1456 gatgcaggaggcctctatcaagtaggtcaattccctcactggaatgtttggtctgagtgg 1515
           |||||||||| |||||||||| |||||| || ||||||||||| |||||||| ||||||
Db    1513 GATGCAGGAGGTCTCTATCAAGAAGGTCAGTTTCCTCACTGGAACGTTTGGTCAGAGTGG 1572

Qy    1516 aatgggaagtaccgggacattgtgcgtcaattcattaaaggcactgatggatttgctggt 1575
           ||||| ||||| || || |  |||||||| ||||| |||||||| ||||||||||||||||
Db    1573 AATGGAAAGTATCGCGATACCGTGCGTCAGTTCATCAAAGGCACAGATGGATTTGCTGGT 1632
```

```
Qy   1576 ggttttgccgaatgtctttgtggaagtccacacctataccaggtaagttgtggcaatact 1635
          | ||||||| ||||| || ||||||||||||||  |||||
Db   1633 GCTTTTGCTGAATGCCTATGTGGAAGTCCACAGTTATAC---------------------- 1671

Qy   1636 tgtaaatgagttgagtgaatgtcacctggatttttttatatataccacatgatgatacaca 1695

Db   1672 ------------------------------------------------------------ 1671

Qy   1696 tctaaatatataacaatcatagtgtatgcatatgcatttggctaagaagtattagtgtat 1755

Db   1672 ------------------------------------------------------------ 1671

Qy   1756 acactagtgctatatataggttttaacacccaacttgccaatgaaggaacatagggcttt 1815

Db   1672 ------------------------------------------------------------ 1671

Qy   1816 ctagttatcttatttatttgtccggtgaataatccactgaaaaattccagccatgtcatt 1875

Db   1672 ------------------------------------------------------------ 1671

Qy   1876 ttttaggggggggagaagaaactatattgatttgccccccctaaaagaagccatctcagaat 1935

Db   1672 ------------------------------------------------------------ 1671

Qy   1936 tcataggtaagttgcttttctgtaaagaaaggaaaacgacttcatactttctatcggtgc 1995

Db   1672 ------------------------------------------------------------ 1671

Qy   1996 taacttagctcgatgtatatttgtaagatgaatgccaaatttaatttgtcggataatttg 2055

Db   1672 ------------------------------------------------------------ 1671

Qy   2056 atctgttattcacaaatttctatttggtttctctagaaatcaaaccagtaacttgttatt 2115

Db   1672 ------------------------------------------------------------ 1671

Qy   2116 ggcactgcaacttcttattgattaatcaggcaggaggaaggaaaccttggcacagtatca 2175
                                 |||||||| || ||||| ||||||||||||||||
Db   1672 -------------------------CAGGCAGGGGGGAGGAAGCCTTGGCACAGTATCG 1705

Qy   2176 actttgtatgtgcacatgatggatttacactggctgatttggtaacatatataagaagt 2235
          |||||||||||||||| ||||||||||||||||||||||||||| ||||| ||||  ||||
Db   1706 GCTTTGTATGTGCACACGATGGATTTACACTGGCTGATTTGGTCACATACAATAGCAAGT 1765

Qy   2236 acaatttaccaaatggggagaacaacagagatggagaaaatcacaatcttagctggaatt 2295
          |||| ||  ||||||| ||| ||  |||||||||  |||||||| ||||||||||||||||
Db   1766 ACAACTTGTCAAATGGTGAGGACTTCAGAGATGGGGAAAATCATAATCTTAGCTGGAATT 1825

Qy   2296 gtggggaggaaggagaattcgcaagattgtctgtcaaaagattgaggaagaggcagatgc 2355
          |||||||||||||||||| |||||  |||| ||| .|||||| |||||||||| ||||
Db   1826 GTGGGGAGGAAGGAGAATTTGCAAGTCTGTCAGTCCGAAGATTAAGGAAGAGGCAAATGC 1885

Qy   2356 gcaatttctttgtttgtctcatggtttctcaaggagttccaatgttctacatgggtgatg 2415
          |||||||||||||||||| |||||||||| ||||||||||||||||||||| ||||
Db   1886 GCAATTTCTTTGTTTGTCTTATGGTTTCTCAGGGAGTTCCAATGTTCTACATGGGCGATG 1945

Qy   2416 aatatggccacacaaaaggggcaacaacaatacatactgccatgattcttatgtcaatt 2475
          ||||||| |||||||| || || |||||||||| |||||||||||   ||||||||||
Db   1946 AATATGGTCACACAAAGGGAGGGAACAACAATACGTACTGCCATGACCATTATGTCAATT 2005

Qy   2476 attttcgctgggataaaa---aagaacaatactctgagttgcaccgattctgctgcctca 2532
          |||| || |||||||| |   |||||||| |||||| ||| || |||||||| | ||||
Db   2006 ATTTCCGTTGGGATAAGAAGGAAGAACAATCCTCTGATTTGTACAGATTCTGCCGTCTCA 2065

Qy   2533 tgaccaaattccgcaaggagtgcgagggtcttggccttgaggactttccaacggccaaac 2592
          ||||| |||||||||| || || ||   |||||||||||||||||| || ||  |  |||
Db   2066 TGACCGAATTCCGCAAAGAATGTGAATCTCTTGGCCTTGAGGACTTCCCGACTTCAGAAC 2125

Qy   2593 ggctgcagtggcatggtcatcagcctgggaagcctgattggtctgagaatagccgattcg 2652
          || || | ||||| |||||||||| |||||||||| ||||| |||   ||||||||||
Db   2126 GGTTGAAATGGCACGGTCATCAGCCCGGGAAGCCTGACTGGTCAGAGGCAAGCCGATTCG 2185

Qy   2653 ttgccttttccatgaaagatgaaagacagggcgagatctatgtggccttcaacaccagcc 2712
          |||||||  |||||||| || ||||   | |||||||||| |||||||||||||||| |
Db   2186 TTGCCTTCACCATGAAGGACGAAACCAAAGGCGAGATCTACGTGGCCTTCAACACCAGTC 2245

Qy   2713 acttaccggccgttgttgagctcccagagcgcgcagggcgccggtgggaaccggtggtgg 2772
          || | ||||  |||||||| ||| ||||||||| | |||  ||| ||||| ||||||||||
Db   2246 ACCTTCCGGTGGTTGTTGGGCTTCCAGAGCGCTCTGGGTTCCGATGGGAGCCGGTGGTGG 2305

Qy   2773 acacaggcaagccagcaccatacgacttcctcaccgacgacttacctgatcgcgctctca 2832
          |||| ||||||   |||||||| |||||||||||||| | | | || ||||| ||| |||
Db   2306 ACACCGGCAAGGAGGCACCATATGACTTCCTCACCGATGGCCTGCCAGATCGTGCTGTCA 2365

Qy   2833 ccatacaccagttctcgcatttcctctactccaacctctaccccatgctcagctactcat 2892
          || |  |||||||||| |||||||||  |||||||  ||||| || ||||||||||||| |
Db   2366 CCGTCTACCAGTTCTCTCATTTCCTCAACTCCAATCTCTATCCTATGCTCAGCTACTCCT 2425

Qy   2893 cggtcatcctagtattgcgccctgatgtttgagagaccaatatata 2938
          |  ||||||| |||||||||||||||| ||| |||   | ||| |
Db   2426 CCATCATCCTTGTATTGCGCCCTGATGTCTGAAAGAAGCAGATACA 2471


RESULT  14
AF142590
LOCUS       AF142590                983 bp    mRNA    linear   PLN 12-AUG-1999
DEFINITION  Triticum aestivum isoamylase 1 mRNA, partial cds.
ACCESSION   AF142590
VERSION     AF142590.1  GI:4929176
KEYWORDS    .
SOURCE      bread wheat.
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Triticum.
REFERENCE   1  (bases 1 to 983)
  AUTHORS   Sun,C., Sathish,P., Ahlandsberg,S. and Jansson,C.
  TITLE     Analyses of isoamylase gene activity in wild-type barley indicate
            its involvement in starch synthesis
  JOURNAL   Plant Mol. Biol. 40 (3), 431-443 (1999)
  MEDLINE   99364538
   PUBMED   10437827
REFERENCE   2  (bases 1 to 983)
  AUTHORS   Sun,C., Sathish,P., Ahlandsberg,S. and Jansson,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-APR-1999) Plant Biology, Genetic Center, The Swedish
            University of Agricultural Sciences, P.O. Box 7080, Uppsala
            S-75007, Sweden
FEATURES             Location/Qualifiers
     source          1. .983
                     /organism="Triticum aestivum"
                     /db_xref="taxon:4565"
     CDS             <1. .>983
                     /codon_start=3
                     /product="isoamylase 1"
                     /protein_id="AAD33890.1"
                     /db_xref="GI:4929177"
                     /translation="DLPLRYPQKDLVIYEMHLRGFTKHDSSNVEHPGTFIGAVSKLDY
                     LKELGVNCIELMPCHEFNELEYSTSSSKMNFWGYSTINFFSPMTRYTSGGIKNCGRDA
                     INEFKTFVREAHKRGIEVILDVVFNHTAEGNENGPILSFRGVDNTTYYMLAPKGEFYN
                     YSGCGNTFNCNHPVVRQFIVDCLRYWVTEMHVDGFRFDLASIMTRGSSLWDPVNVYGA
                     PIEGDMITTGTPLVTPPLIDMISNDPILGGVKLVAEAWDAGGLYQVGQFPHWNVWSEW
                     NGKYRDIVRQFIKGTDGFAGGFAECLCGSPHLYQAGGRKPWHSINFVCAHD"
BASE COUNT      271 a    188 c    241 g    283 t
ORIGIN


  Query Match            31.1%;  Score 932.8;  DB 8;  Length 983;
  Best Local Similarity  98.7%;  Pred. No. 5e-165;
  Matches  940;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy    686 gcgacctacctctaagatatcctcaaaaggacctggtaatatatgagatgcacttgcgtg 745
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GCGACCTACCTCTAAGATATCCTCAAAAGGACCTGGTAATATATGAGATGCACTTGCGTG 60

Qy    746 gattcacgaagcatgattcaagcaatgtagaacatccgggtactttcattggagctgtgt 805
```

```
Db     61 GATTCACGAAGCATGATTCAAGCAATGTAGAACATCCGGGTACTTTCATTGGAGCTGTGT 120

Qy    806 cgaagcttgactatttgaaggagcttggagttaattgtattgaattaatgccctgccatg 865
Db    121 CGAAGCTTGACTATTTGAAGGAGCTTGGAGTTAATTGTATTGAATTAATGCCCTGCCATG 180

Qy    866 agttcaacgagctggagtactcaacctcttcttccaagatgaacttttggggatattcta 925
Db    181 AGTTCAACGAGCTGGAGTACTCAACCTCTTCTTCCAAGATGAACTTTTGGGGATATTCTA 240

Qy    926 ccataaacttcttttcaccaatgacaagatacacatcaggcgggataaaaaactgtgggc 985
Db    241 CCATAAACTTCTTTTCACCAATGACGAGATACACATCAGGCGGGATAAAAAACTGTGGGC 300

Qy    986 gtgatgccataaatgagttcaaaacttttgtaagagaggctcacaaacggggaattgagg 1045
Db    301 GTGATGCCATAAATGAGTTCAAAACTTTTGTAAGAGAGGCTCACAAACGGGGAATTGAGG 360

Qy   1046 tgatcctggatgttgtcttcaaccatacagctgagggtaatgagaatggtccaatattat 1105
Db    361 TGATCCTGGATGTTGTCTTCAACCATACAGCTGAGGGTAATGAGAATGGTCCAATATTAT 420

Qy   1106 catttaaggggtcgataatactacatactatatgcttgcacccaagggagagttttata 1165
Db    421 CATTTAGGGGGGTCGATAATACTACATACTATATGCTTGCACCCAAGGGAGAGTTTTATA 480

Qy   1166 actattctggctgtgggaataccttcaactgtaatcatcctgtggttcgtcaattcattg 1225
Db    481 ACTATTCTGGCTGTGGGAATACCTTCAACTGTAATCATCCTGTGGTTCGTCAATTCATTG 540

Qy   1226 tagattgtttaagatactgggtgacggaaatgcatgttgatggttttcgttttgatcttg 1285
Db    541 TAGATTGTTTAAGATACTGGGTGACGGAAATGCATGTTGATGGTTTTCGTTTTGATCTTG 600

Qy   1286 catccataatgaccagaggttccagtctgtgggatccagttaacgtgtatggagctccaa 1345
Db    601 CATCCATAATGACCAGAGGTTCCAGTCTGTGGGATCCAGTTAACGTGTATGGAGCTCCAA 660

Qy   1346 tagaaggtgacatgatcacaacagggacacctcttgttactccaccacttattgacatga 1405
Db    661 TAGAAGGTGACATGATCACAACAGGGACACCTCTTGTTACTCCACCACTTATTGACATGA 720

Qy   1406 tcagcaatgacccaattcttggaggcgtcaagctcattgctgaagcatgggatgcaggag 1465
Db    721 TCAGCAATGACCCAATTCTTGGAGGCGTCAAGCTCGTTGCTGAAGCATGGGATGCAGGAG 780

Qy   1466 gcctctatcaagtaggtcaattccctcactggaatgtttggtctgagtggaatgggaagt 1525
Db    781 GCCTCTATCAAGTAGGTCAATTCCCTCACTGGAATGTTTGGTCTGAGTGGAATGGGAAGT 840

Qy   1526 accgggacattgtgcgtcaattcattaaaggcactgatggatttgctggtggttttgccg 1585
Db    841 ACCGGGACATTGTGCGTCAATTCATTAAAGGCACTGATGGATTTGCTGGTGGTTTTGCCG 900

Qy   1586 aatgtctttgtggaagtccacacctataccaggtaagttgtggcaatacttg 1637
Db    901 AATGTCTTTGTGGAAGTCCACACCTATACCAGGCAGGAGGAAGGAAACCTTG 952


RESULT  15
AX006091
LOCUS       AX006091               2706 bp    DNA     linear   PAT 24-AUG-2000
DEFINITION  Sequence 2 from Patent WO9906575.
ACCESSION   AX006091
VERSION     AX006091.1  GI:9929017
KEYWORDS    .
SOURCE      potato.
  ORGANISM  Solanum tuberosum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
REFERENCE   1  (bases 1 to 2706)
  AUTHORS   Martin,C.R. and Smith,A.M.
  TITLE     Starch debranching enzymes
  JOURNAL   Patent: WO 9906575-A 2 11-FEB-1999;
            MARTIN CATHERINE ROSEMARY (GB); SMITH ALISON MARY (GB)
FEATURES             Location/Qualifiers
     source          1. .2706
                     /organism="Solanum tuberosum"
                     /db_xref="taxon:4113"
     misc_feature    2641
                     /note="n is a or g or c or t"
BASE COUNT      797 a    492 c    624 g    792 t      1 others
ORIGIN


  Query Match             22.5%;  Score 673.2;  DB 6;  Length 2706;
  Best Local Similarity   69.4%;  Pred. No. 3.2e-116;
  Matches  937;  Conservative    0;  Mismatches  398;  Indels   15;  Gaps    1;

Qy    395 aggataggg tgaccgaggaggttccccttgacccccgatgaatcggaccgggaacgtgt 454
Db    425 AGAAGAGAGTGACCGAGCAAATTTTCCTGGATCCTCTAGCTAATAAAACTGGAGATGTAT 484

Qy    455 ggcatgtcttcatcgaaggcgagctgcacaacatgctttacgggtacaggttcgacggca 514
Db    485 GGCATGTGTTCCTTAAGGGAGATTTTGAGAATATGCTATATGGCTACAAATTTGATGGGA 544

Qy    515 cctttgctcctcactgcgggcactaccttgatgtttccaatgtcgtggtggatccttatg 574
Db    545 AATTCTGTCCTGAAGAAGGACACTACTTTGACTCTTCGCAGATAGTGTTGGATCCTTATG 604

Qy    575 ctaaggcagtgataagccgaggggagtatggtgttccagcgcgtggtaacaattgctggc 634
Db    605 CCAAGGCTATAGTAAGCAGAGGAGAATATGGTGTATTAGGGCCAGAGGATGATTGTTGGC 664

Qy    635 ctcagatggctggcatgatccctcttccatatagcacgtttgattgggaaggcgacctac 694
Db    665 CCCCAATGGCTGGCATGGTACCTTCTGCTTCTGATCAGTTTGATTGGGAAGGAGATCTAC 724

Qy    695 ctctaagatatcctcaaaaggacctggtaatatatgagatgcacttgcgtggattcacga 754
Db    725 CACTGAAGTTTCCACAGAGAGATCTTGTAATCTATGAAATGCATGTTCGTGGGTTTACTA 784

Qy    755 agcatgattcaagcaatgtagaacatccgggtactttcattggagctgtgtcgaagcttg 814
Db    785 ATCATGAGTCGAGTGAAACAAAATATCCTGGTACTTACCTTGGTGTTGTGGAGAAACTTG 844

Qy    815 actatttgaaggagcttggagttaattgtattgaattaatgccctgccatgagttcaacg 874
Db    845 ATCACTTGAAGGAACTTGGTGTCAACTGTATAGAGCTAATGCCCTGTCACGAGTTCAATG 904

Qy    875 agctggagtactcaacctcttcttc----------caagatgaacttttggggat 919
Db    905 AGCTGGAGTACTATAGTTATAACTCTGTATTGGGCGACTACAAGTTTAACTTTTGGGGCT 964

Qy    920 attctaccataaacttcttttcaccaatgacaagatacacatcaggcgggataaaaaact 979
Db    965 ATTCTACTGTCAATTTCTTTTCTCCAATGGGAAGATACTCATCTGCTGGTCTAAGTAATT 1024

Qy    980 gtgggcgtgatgccataaatgagttcaaaacttttgtaagagaggctcacaaacggggaa 1039
Db   1025 GCGGCCTCGGTGCAATAAACGAATTTAAGTATCTTGTCAAGGAAGCACATAAACGTGGAA 1084

Qy   1040 ttgaggtgatcctggatgttgtcttcaaccatacagctgagggtaatgagaatggtccaa 1099
Db   1085 TCGAGGTTATCATGGATGTTGTTTTCAATCACACTGCTGAAGGAAATGAAAATGGTCCCA 1144

Qy   1100 tattatcatttaaggggtcgataatactacatactatatgcttgcacccaagggagagt 1159
Db   1145 TACTATCATTTAGAGGCATTGACAACAGTGTGTTTTATACGCTAGCTCCTAAGGGTGAAT 1204

Qy   1160 tttataactattctggctgtgggaataccttcaactgtaatcatcctgtggttcgtcaat 1219
Db   1205 TTTACAACTACTCAGGATGTGGAAATACCTTCAACTGTAATAATCCCATTGTACGTCAAT 1264
```

```
Qy   1220 tcattgtagattgtttaagatactgggtgacggaaatgcatgttgatggttttcgttttg 1279
          | || || ||||| || ||||| ||||| || |||||||| || ||||| || || ||||
Db   1265 TTATAGTGGATTGCTTGAGATATTGGGTTACCGAAATGCACGTAGATGGCTTCCGCTTTG 1324

Qy   1280 atcttgcatccataatgaccagaggttccagtctgtgggatccagttaacgtgtatggag 1339
          ||||||| || ||  | || ||| ||  |||   |||| || | || || || ||||||
Db   1325 ATCTTGCTTCTATCCTTACAAGAAGTAGCAGCTCGTGGAATGCTGTAAATGTCTATGGAA 1384

Qy   1340 ctccaatagaaggtgacatgatcacaacagggacacctcttgttactccaccacttattg 1399
           | |||| || |||||| ||||||| ||||| || |||||    |  |||||| | ||||
Db   1385 ATTCAATTGACGGTGACGTGATCACCACAGGCACTCCTCTCACAAGCCCACCATTGATTG 1444

Qy   1400 acatgatcagcaatgacccaattcttggaggcgtcaagctcattgctgaagcatgggatg 1459
          | ||||| |||||||| ||||| ||| | || || ||||| || ||||||||||||||
Db   1445 ATATGATTAGCAATGATCCAATACTTCGTGGAGTAAAGCTTATAGCTGAAGCATGGGATT 1504

Qy   1460 caggaggcctctatcaagtaggtcaattccctcactggaatgtttggtctgagtggaatg 1519
            |||||||| || ||||| ||    || || ||||||  | | ||||| ||||||||
Db   1505 GTGGAGGCCTTTACCAAGTTGGCATGTTTCCGCACTGGGGTATCTGGTCGGAGTGGAACG 1564

Qy   1520 ggaagtaccgggacattgtgcgtcaattcattaaaggcactgatggatttgctggtggtt 1579
          | |||||||| ||||| || || || ||||| ||||||||||||| ||| |||| | ||
Db   1565 GAAAGTACCGTGACATGGTACGGCAGTTCATCAAAGGCACTGATGGGTTTTCTGGGGCTT 1624

Qy   1580 ttgccgaatgtctttgtggaagtccacacctataccaggtaagttgtggcaatacttgta 1639
          |||| ||||| |||||||||| ||| | |||||||||  | |  |  | ||  | || |
Db   1625 TTGCTGAATGCCTTTGTGGAAGCCCAAATCTATACCAGAAAGGAGGAAGAAAACCATGGA 1684

Qy   1640 aatgagttgagtgaatgtcacctggatttttatatataccacatgatgatacacatcta 1699
          |  |  |  | |   |||   |       |  | |   |    |    | |   |||| |
Db   1685 ACAGTATAAATTTCGTGTGTGCCCACGATGGTTTTACTTTGGCTGATTTAGTGACATACA 1744

Qy   1700 aatatataacaatcatagtgtatgcatatg 1729
          |  ||| |   |   | |   ||| | | |
Db   1745 ACAATAAACACAATTTGGCAAATGGAGAGG 1774
```

Search completed: July 31, 2002, 10:14:30
Job time: 7149 sec

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: July 31, 2002, 09:08:46 ; Search time 318.82 Seconds
(without alignments)
16139.491 Million cell updates/sec

Title: US-09-674-817A-1
Perfect score: 2997
Sequence: 1 ggtcggggccggcgccgcgc.....aaaaaaaaaaaaaaaaaaaa 2997

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802:*
1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	2997	100.0	2997	21	AAZ24492	Wheat isoamylase c
	2	1074.4	35.8	1706	20	AAX34654	Partial cDNA for h
	3	1000	33.4	2712	20	AAX76375	Zea mays SU1 starc
	4	673.2	22.5	2706	20	AAX27061	S. tuberosum isoam
	5	613.2	20.5	2806	20	AAX27066	Original S. tubero
	6	575.4	19.2	2133	18	AAT96754	S. tuberosum debra
	7	533.4	17.8	9289	20	AAX34655	Hexaploid wheat DB
	8	405.8	13.5	1663	20	AAX33162	Potato isoamylase
	9	149.2	5.0	2634	20	AAX27062	S. tuberosum isoam
c	10	143	4.8	273254	21	AAC81914	Chlamydia pneumoni
	11	143	4.8	1230025	20	AAX91990	Nucleotide sequenc
	12	129.6	4.3	1389	20	AAX27067	Original S. tubero
	13	120.8	4.0	1038602	20	AAZ01425	Complete genome se
	14	120	4.0	2523	20	AAX76376	Zea mays partial s
	15	116.2	3.9	22934	23	AAS59613	Propionibacterium
	16	103.8	3.5	1496	22	AAF71542	Corynebacterium gl
	17	103.8	3.5	1913	22	AAF71541	Corynebacterium gl
	18	103.8	3.5	2508	22	AAH67270	C glutamicum codin
	19	103.8	3.5	349980	22	AAH68531	C glutamicum codin
c	20	102.8	3.4	4411529	22	AAI99682	Mycobacterium tube
	21	101.8	3.4	1965	23	AAS56077	Salmonella typhi D
	22	83.8	2.8	1557	23	AAS88837	DNA encoding novel
	23	83.8	2.8	3444	23	AAS88835	DNA encoding novel
c	24	83.8	2.8	3912	23	AAS89192	DNA encoding novel
	25	80.6	2.7	2181	20	AAX07446	Rhodothermus marin
	26	80.6	2.7	2181	21	AAZ89961	Isoamylase nucleot
	27	70.4	2.3	1215	21	AAA63296	C. trachomatis LGV
	28	70.4	2.3	1215	22	AAH56200	Chlamydia trachoma
	29	66	2.2	3467	17	AAT12324	DNA encoding Sulfo
	30	63.2	2.1	463	20	AAX33154	Potato isoamylase
	31	63.2	2.1	501	20	AAX33158	Potato isoamylase
	32	63.2	2.1	2109	20	AAZ07194	Human lung tumour
	33	63.2	2.1	2109	21	AAC79147	Human lung tumour-
	34	63.2	2.1	2109	23	AAD23222	Human lung tumour-
	35	63.2	2.1	13862	22	ABA08208	Human ovarian and
	36	63.2	2.1	13862	22	AAL02789	Human reproductive
	37	63.2	2.1	13862	22	AAL07516	Human reproductive
	38	61.8	2.1	468	20	AAX33155	Potato isoamylase
	39	61.8	2.1	503	20	AAX33160	Potato isoamylase
c	40	61.2	2.0	117213	19	AAV62176	HSV-2 strain SB5 C
c	41	59	2.0	71989	21	AAA29349	Sorangium cellulos
	42	58.4	1.9	408	21	AAA67250	Pinus radiata amyl
	43	57.2	1.9	1631	22	AAH33111	Human colon cancer
	44	57	1.9	154746	24	AAD25519	Human herpesvirus
c	45	57	1.9	154746	24	AAD25519	Human herpesvirus

ALIGNMENTS

RESULT 1
AAZ24492
ID AAZ24492 standard; DNA; 2997 BP.
XX
AC AAZ24492;
XX
DT 18-FEB-2000 (first entry)
XX
DE Wheat isoamylase cDNA.
XX
KW Isoamylase; wheat; transgenic plant; starch production; food; baking;
KW pastry; packaging material; glucose; glucan; paper; pulp; adhesive;
KW textile building material; soil stabilizer; wetting agent; fertilizer;
KW plant-protection; cosmetic; flocculant; ss.
XX
OS Triticum aestivum.
XX
FH Key Location/Qualifiers
FT CDS 1..2924
FT /*tag= a
FT /product= "isoamylase"
FT /codon_start= 3
FT /note= "coding sequence is interrupted by introns
FT and the start codon is not disclosed"
FT exon 1..295
FT /*tag= b
FT /number= 1
FT intron 296..395
FT /*tag= c
FT /number= 1
FT exon 396..1616
FT /*tag= d

```
FT                   /number= 2
FT   intron          1617..2134
FT                   /*tag=  e
FT                   /number= 2
FT   exon            2135..2924
FT                   /*tag=  f
FT                   /number= 3
XX
PN   DE19820608-A1.
XX
PD   11-NOV-1999.
XX
PF   08-MAY-1998;   98DE-1020608.
XX
PR   08-MAY-1998;   98DE-1020608.
XX
PA   (AGRE ) HOECHST-SCHERING AGREVO GMBH.
XX
PI   Loerz H,  Luetticke S,  Abel G;
XX
DR   WPI; 2000-024509/03.
DR   P-PSDB; AAY50819.
XX
PT   New nucleic acid encoding isoamylase from wheat and related transgenic
PT   plants producing starch with altered properties  -
XX
PS   Claim 1b; Page 14-15; 18pp; German.
XX
CC   This invention describes a novel wheat isoamylase. Transgenic plants,
CC   specifically wheat, that contain (I) are used for production of starch,
CC   used particularly in foods, particularly baked and pastry goods and for
CC   making packaging materials or disposable items. Starch may also be used
CC   as starting materials for glucose or glucan components (e.g. for
CC   fermentation or further chemical conversion); in paper and pulp
CC   production; as adhesives; in textiles; in preparation of gypsum-based
CC   building materials; as soil stabilizer; as wetting agent etc. in
CC   fertilizer and plant-protection compositions; as binder (in
CC   pharmaceuticals, cosmetics, coal briquetting and casting sand); as
CC   flocculant in soil or coal slurries; as rubber and leather additives; and
CC   for production of synthetic polymers, e.g. polyurethane films. Transgenic
CC   plants with increased/decreased production of (I) produce starches with
CC   altered physical and/or chemical properties, e.g. amylose/amylopectin
CC   ratios; degree of branching; mean chain length; phosphate content;
CC   gelatinization properties; gel- or film-forming properties, or starch
CC   grain size or structure, e.g. easier, and less expensive, digestion by
CC   enzymes. This sequence encodes the wheat isoamylase protein described in
CC   the invention.
XX
SQ   Sequence 2997 BP; 803 A; 636 C; 762 G; 796 T; 0 other;

  Query Match             100.0%;  Score 2997;  DB 21;  Length 2997;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 2997;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ggtcggggccggcgccgcgcctgcgacggtggcgacccaatgcgacggcggggaaggggg 60
Db       1 ggtcggggccggcgccgcgcctgcgacggtggcgacccaatgcgacggcggggaaggggg 60

Qy      61 tcggcgaggtgtgcgccgcggttgtcgaggcggcgacgaaggtagaggacgaggggggagg 120
Db      61 tcggcgaggtgtgcgccgcggttgtcgaggcggcgacgaaggtagaggacgaggggggagg 120

Qy     121 aggacgagccggtggcggaggacaggtacgcgctcggcggcgcgtgcagggtgctcgccg 180
Db     121 aggacgagccggtggcggaggacaggtacgcgctcggcggcgcgtgcagggtgctcgccg 180

Qy     181 gaatgcccgcgccgctgggcgccaccgcgctcgccggcggggtcaatttcgccgtctatt 240
Db     181 gaatgcccgcgccgctgggcgccaccgcgctcgccggcggggtcaatttcgccgtctatt 240

Qy     241 ccggcggagccaccgccgcggcgctctgcctcttcacgccagaagatctcaaggcggtgg 300
Db     241 ccggcggagccaccgccgcggcgctctgcctcttcacgccagaagatctcaaggcggtgg 300

Qy     301 ggttgcctcccgagtagagttcatcagctttgcgtgcgccgcgcgccccttttttgggcc 360
Db     301 ggttgcctcccgagtagagttcatcagctttgcgtgcgccgcgcgccccttttttgggcc 360

Qy     361 tgcaatttaagttttgtactggggcaaatgctgcaggatagggtgaccgaggaggttccc 420
Db     361 tgcaatttaagttttgtactggggcaaatgctgcaggatagggtgaccgaggaggttccc 420

Qy     421 cttgacccccctgatgaatcggaccgggaacgtgtggcatgtcttcatcgaaggcgagctg 480
Db     421 cttgacccccctgatgaatcggaccgggaacgtgtggcatgtcttcatcgaaggcgagctg 480

Qy     481 cacaacatgctttacgggtacaggttcgacggcacctttgctcctcactgcgggcactac 540
Db     481 cacaacatgctttacgggtacaggttcgacggcacctttgctcctcactgcgggcactac 540

Qy     541 cttgatgtttccaatgtcgtggtggatccttatgctaaggcagtgataagccgaggggag 600
Db     541 cttgatgtttccaatgtcgtggtggatccttatgctaaggcagtgataagccgaggggag 600

Qy     601 tatggtgttccagcgcgtggtaacaattgctggcctcagatggctggcatgatccctctt 660
Db     601 tatggtgttccagcgcgtggtaacaattgctggcctcagatggctggcatgatccctctt 660

Qy     661 ccatatagcacgtttgattgggaaggcgacctacctctaagatatcctcaaaaggacctg 720
Db     661 ccatatagcacgtttgattgggaaggcgacctacctctaagatatcctcaaaaggacctg 720

Qy     721 gtaatatatgagatgcacttgcgtggattcacgaagcatgattcaagcaatgtagaacat 780
Db     721 gtaatatatgagatgcacttgcgtggattcacgaagcatgattcaagcaatgtagaacat 780

Qy     781 ccgggtactttcattggagctgtgtcgaagcttgactatttgaaggagcttggagttaat 840
Db     781 ccgggtactttcattggagctgtgtcgaagcttgactatttgaaggagcttggagttaat 840

Qy     841 tgtattgaattaatgccctgccatgagttcaacgagctggagtactcaacctcttcttcc 900
Db     841 tgtattgaattaatgccctgccatgagttcaacgagctggagtactcaacctcttcttcc 900

Qy     901 aagatgaacttttggggatattctaccataaacttcttttcaccaatgacaagatacaca 960
Db     901 aagatgaacttttggggatattctaccataaacttcttttcaccaatgacaagatacaca 960

Qy     961 tcaggcgggataaaaaactgtgggcgtgatgccataaatgagttcaaaacttttgtaaga 1020
Db     961 tcaggcgggataaaaaactgtgggcgtgatgccataaatgagttcaaaacttttgtaaga 1020

Qy    1021 gaggctcacaaacggggaattgaggtgatcctggatgttgtcttcaaccatacagctgag 1080
Db    1021 gaggctcacaaacggggaattgaggtgatcctggatgttgtcttcaaccatacagctgag 1080

Qy    1081 ggtaatgagaatggtccaatattatcatttaaggggtcgataatactacatactatatg 1140
Db    1081 ggtaatgagaatggtccaatattatcatttaaggggtcgataatactacatactatatg 1140

Qy    1141 cttgcacccaagggagagttttataactattctggctgtgggaatacctcaactgtaat 1200
Db    1141 cttgcacccaagggagagttttataactattctggctgtgggaatacctcaactgtaat 1200

Qy    1201 catcctgtggttcgtcaattcattgtagattgtttaagatactgggtgacggaaatgcat 1260
Db    1201 catcctgtggttcgtcaattcattgtagattgtttaagatactgggtgacggaaatgcat 1260

Qy    1261 gttgatggttttcgttttgatcttgcatccataatgaccagaggttccagtctgtgggat 1320
Db    1261 gttgatggttttcgttttgatcttgcatccataatgaccagaggttccagtctgtgggat 1320

Qy    1321 ccagttaacgtgtatggagctccaatagaaggtgacatgatcacaacagggacacctctt 1380
Db    1321 ccagttaacgtgtatggagctccaatagaaggtgacatgatcacaacagggacacctctt 1380
```

```
Qy   1381 gttactccaccacttattgacatgatcagcaatgacccaattcttggaggcgtcaagctc 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 gttactccaccacttattgacatgatcagcaatgacccaattcttggaggcgtcaagctc 1440

Qy   1441 attgctgaagcatgggatgcaggaggcctctatcaagtaggtcaattccctcactggaat 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 attgctgaagcatgggatgcaggaggcctctatcaagtaggtcaattccctcactggaat 1500

Qy   1501 gtttggtctgagtggaatgggaagtaccgggacattgtgcgtcaattcattaaaggcact 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 gtttggtctgagtggaatgggaagtaccgggacattgtgcgtcaattcattaaaggcact 1560

Qy   1561 gatggatttgctggtggttttgccgaatgtctttgtggaagtccacacctataccaggta 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 gatggatttgctggtggttttgccgaatgtctttgtggaagtccacacctataccaggta 1620

Qy   1621 agttgtggcaatacttgtaaatgagttgagtgaatgtcacctggatttttttatatatacc 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 agttgtggcaatacttgtaaatgagttgagtgaatgtcacctggatttttttatatatacc 1680

Qy   1681 acatgatgatacacatctaaatatataacaatcatagtgtatgcatatgcatttggctaa 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 acatgatgatacacatctaaatatataacaatcatagtgtatgcatatgcatttggctaa 1740

Qy   1741 gaagtattagtgtatacactagtgctatatataggttttaacacccaacttgccaatgaa 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 gaagtattagtgtatacactagtgctatatataggttttaacacccaacttgccaatgaa 1800

Qy   1801 ggaacatagggctttctagttatcttatttatttgtccggtgaataatccactgaaaaat 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 ggaacatagggctttctagttatcttatttatttgtccggtgaataatccactgaaaaat 1860

Qy   1861 tccagccatgtcatttttttaggggggagaagaaactatattgatttgccccctaaaag 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 tccagccatgtcatttttttaggggggagaagaaactatattgatttgccccctaaaag 1920

Qy   1921 aagccatctcagaattcataggtaagttgctttttctgtaaagaaaggaaaacgacttcat 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 aagccatctcagaattcataggtaagttgctttttctgtaaagaaaggaaaacgacttcat 1980

Qy   1981 actttctatcggtgctaacttagctcgatgtatatttgtaagatgaatgccaaatttaat 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 actttctatcggtgctaacttagctcgatgtatatttgtaagatgaatgccaaatttaat 2040

Qy   2041 ttgtcggataatttgatctgttattcacaaatttctatttggtttctctagaaatcaaac 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 ttgtcggataatttgatctgttattcacaaatttctatttggtttctctagaaatcaaac 2100

Qy   2101 cagtaacttgttattggcactgcaacttcttattgattaatcaggcaggaggaaggaaac 2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 cagtaacttgttattggcactgcaacttcttattgattaatcaggcaggaggaaggaaac 2160

Qy   2161 cttggcacagtatcaactttgtatgtgcacatgatggatttacactggctgatttggtaa 2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 cttggcacagtatcaactttgtatgtgcacatgatggatttacactggctgatttggtaa 2220

Qy   2221 catataataagaagtacaatttaccaaatggggagaacaacagagatggagaaaatcaca 2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 catataataagaagtacaatttaccaaatggggagaacaacagagatggagaaaatcaca 2280

Qy   2281 atcttagctggaattgtggggaggaaggagaattcgcaagattgtctgtcaaaagattga 2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 atcttagctggaattgtggggaggaaggagaattcgcaagattgtctgtcaaaagattga 2340

Qy   2341 ggaagaggcagatgcgcaatttctttgtttgtctcatggtttctcaaggagttccaatgt 2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 ggaagaggcagatgcgcaatttctttgtttgtctcatggtttctcaaggagttccaatgt 2400

Qy   2401 tctacatgggtgatgaatatggccacacaaaaggggcaacaacaatacatactgccatg 2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401 tctacatgggtgatgaatatggccacacaaaaggggcaacaacaatacatactgccatg 2460
```

```
Qy   2461 attcttatgtcaattattttcgctgggataaaaaaagaacaatactctgagttgcaccgat 2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2461 attcttatgtcaattattttcgctgggataaaaaaagaacaatactctgagttgcaccgat 2520

Qy   2521 tctgctgcctcatgaccaaattccgcaaggagtgcgagggtcttggccttgaggactttc 2580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2521 tctgctgcctcatgaccaaattccgcaaggagtgcgagggtcttggccttgaggactttc 2580

Qy   2581 caacggccaaacggctgcagtggcatggtcatcagcctgggaagcctgattggtctgaga 2640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2581 caacggccaaacggctgcagtggcatggtcatcagcctgggaagcctgattggtctgaga 2640

Qy   2641 atagccgattcgttgccttttccatgaaagatgaaagacagggcgagatctatgtggcct 2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2641 atagccgattcgttgccttttccatgaaagatgaaagacagggcgagatctatgtggcct 2700

Qy   2701 tcaacaccagccacttaccggccgttgttgagctcccagagcgcgcagggcgccggtggg 2760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2701 tcaacaccagccacttaccggccgttgttgagctcccagagcgcgcagggcgccggtggg 2760

Qy   2761 aaccggtggtggacacaggcaagccagcaccatacgacttcctcaccgacgacttacctg 2820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2761 aaccggtggtggacacaggcaagccagcaccatacgacttcctcaccgacgacttacctg 2820

Qy   2821 atcgcgctctcaccatacaccagttctcgcatttcctctactccaacctctaccccatgc 2880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2821 atcgcgctctcaccatacaccagttctcgcatttcctctactccaacctctaccccatgc 2880

Qy   2881 tcagctactcatcggtcatcctagtattgcgccctgatgtttgagagaccaatatatataca 2940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2881 tcagctactcatcggtcatcctagtattgcgccctgatgtttgagagaccaatatatataca 2940

Qy   2941 gtaaataatatgtctatatgtaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 2997
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2941 gtaaataatatgtctatatgtaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 2997
```

```
RESULT   2
AAX34654
ID   AAX34654 standard; cDNA; 1706 BP.
XX
AC   AAX34654;
XX
DT   05-JUL-1999  (first entry)
XX
DE   Partial cDNA for hexaploid wheat DBE.
XX
KW   Starch biosynthetic pathway; cereal plant; enzyme; SBE; SSS; DBE; GBSS;
KW   starch branching enzyme; starch soluble synthase; debranching enzyme;
KW   endosperm; wheat; barley; granule-bound synthase; glutenin; starch;
KW   grain softness protein I; bacterial isoamylase; glycogen synthase;
KW   wSBE I-D4 gene; ss.
XX
OS   Triticum tauschii.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1704
FT                   /*tag= a
FT                   /note= "partial DBE; the atrt and stop codons are
FT                         not indicated"
XX
PN   WO9914314-A1.
XX
PD   25-MAR-1999.
XX
PF   11-SEP-1998;   98WO-AU00743.
XX
PR   20-MAR-1998;   98AU-0002509.
PR   12-SEP-1997;   97AU-0009108.
XX
PA   (CSIR ) COMMONWEALTH SCI & IND RES ORG.
PA   (GOOD-) GOODMAN FIELDER LTD.
PA   (LIMA-) GRP LIMAGRAIN PACIFIC PTY LTD.
```

PA (AUSU) UNIV AUSTRALIAN NAT.
XX
PI Li Z, Morell M, Rahman S;
XX
DR WPI; 1999-229525/19.
DR P-PSDB; AAY06918.
XX
PT New isolated cereal plant enzyme genes used for, e.g. expression of
PT antisense sequences of granule bound synthase
XX
PS Claim 47; Page 100-102; 171pp; English.
XX
CC The invention relates to a novel enzyme of starch biosynthetic pathway
CC in a cereal plant, where the enzyme is selected from starch branching
CC enzyme (SBE) I , SBE II, starch soluble synthase (SSS) I, and debranching
CC enzyme (DBE), with the proviso that the enzyme is not SSS I of rice, or
CC SBE I of rice or maize. The methods and products can be used for
CC targeting expression specifically to the endosperm of the seeds of cereal
CC plants such as wheat or barley. They can be used for the expression of
CC e.g. antisense sequences of granule-bound synthase (GBSS), SBE II, low
CC mol. wt. glutenin, grain softness protein I, bacterial isoamylase,
CC bacterial glycogen synthase, and wheat high mol. wt. glutenin Bx17. They
CC can be used for modifying the characteristics of starch produced by a
CC plant. The present sequence represents the wheat DBE partial cDNA
CC sequence.
XX
SQ Sequence 1706 BP; 467 A; 355 C; 415 G; 469 T; 0 other;

Query Match 35.8%; Score 1074.4; DB 20; Length 1706;
Best Local Similarity 74.9%; Pred. No. 4e-222;
Matches 1621; Conservative 0; Mismatches 16; Indels 527; Gaps 1;

```
Qy    799 gctgtgtcgaagcttgactatttgaaggagcttggagttaattgtattgaattaatgccc 858
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 gctgtgtcgaagcttgactatttgaaggagcttggagttaattgtattgaattaatgccc 60

Qy    859 tgccatgagttcaacgagctggagtactcaacctcttcttccaagatgaacttttgggga 918
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 tgccatgagttcaacgagctggagtactcaacctcttcttccaagatgaacttttgggga 120

Qy    919 tattctaccataaacttcttttcaccaatgacaagatacacatcaggcgggataaaaaac 978
          |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db    121 tattctaccataaacttcttttcaccaatgacgagatacacatcaggcgggataaaaaac 180

Qy    979 tgtgggcgtgatgccataaatgagttcaaaacttttgtaagagaggctcacaaacgggga 1038
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 tgtgggcgtgatgccataaatgagttcaaaacttttgtaagagaggctcacaaacgggga 240

Qy   1039 attgaggtgatcctggatgttgtcttcaaccatacagctgagggtaatgagaatggtcca 1098
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 attgaggtgatcctggatgttgtcttcaaccatacagctgagggtaatgagaatggtcca 300

Qy   1099 atattatcatttaaggggtcgataatactacatactatatgcttgcacccaagggagag 1158
          |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db    301 atattatcatttaggggggtcgataatactacatactatatgcttgcacccaagggagag 360

Qy   1159 ttttataactattctggctgtgggaataccttcaactgtaatcatcctgtggttcgtcaa 1218
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 ttttataactattctggctgtgggaatccttcaactgtaatcatcctgtggttcgtcaa 420

Qy   1219 ttcattgtagattgtttaagatactgggtgacggaaatgcatgttgatggttttcgtttt 1278
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db    421 ttcattgtagattgtttaagatactgggtgatggaaatgcatgttgatggttttcgtttt 480

Qy   1279 gatcttgcatccataatgaccagaggttccagtctgtgggatccagttaacgtgtatgga 1338
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 gatcttgcatccataatgaccagaggttccagtctgtgggatccagttaacgtgtatgga 540

Qy   1339 gctccaatagaaggtgacatgatcacaacagggacacctcttgttactccaccacttatt 1398
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 gctccaatagaaggtgacatgatcacaacagggacacctcttgttactccaccacttatt 600

Qy   1399 gacatgatcagcaatgacccaattcttggaggcgtcaagctcattgctgaagcatgggat 1458
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 gacatgatcagcaatgacccaattcttggaggcgtcaagctcattgctgaagcatgggat 660

Qy   1459 gcaggaggcctctatcaagtaggtcaattccctcactggaatgtttggtctgagtggaat 1518
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 gcaggaggcctctatcaagtaggtcaattccctcactggaatgtttggtctgagtggaat 720

Qy   1519 gggaagtaccgggacattgtgcgtcaattcattaaaggcactgatggatttgctggtggt 1578
          |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db    721 gggaagtaccgggacattgtgcgccaattcattaaaggcactgatggatttgctggtggt 780

Qy   1579 tttgccgaatgtctttgtggaagtccacacctataccaggtaagttgtggcaatacttgt 1638
          ||||||||||||||||||||||||||||||||||
Db    781 tttgccgaatgtctttgtggaagtccacacctatac---------------------- 816

Qy   1639 aaatgagttgagtgaatgtcacctggatttttttatatataccacatgatgatacacatct 1698

Db    817 ---------------------------------------------------------- 816

Qy   1699 aaatatataacaatcatagtgtatgcatatgcatttggctaagaagtattagtgtataca 1758

Db    817 ---------------------------------------------------------- 816

Qy   1759 ctagtgctatatataggttttaacacccaacttgccaatgaaggaacatagggctttcta 1818

Db    817 ---------------------------------------------------------- 816

Qy   1819 gttatcttatttatttgtccggtgaataatccactgaaaaattccagccatgtcatttttt 1878

Db    817 ---------------------------------------------------------- 816

Qy   1879 taggggggggagaagaaactatattgatttgccccccctaaaagaagccatctcagaattca 1938

Db    817 ---------------------------------------------------------- 816

Qy   1939 taggtaagttgcttttctgtaaagaaaggaaaacgacttcatactttctatcggtgctaa 1998

Db    817 ---------------------------------------------------------- 816

Qy   1999 cttagctcgatgtatatttgtaagatgaatgccaaatttaatttgtcggataatttgatc 2058

Db    817 ---------------------------------------------------------- 816

Qy   2059 tgttattcacaaatttctatttggtttctctagaaatcaaaccagtaacttgttattggc 2118

Db    817 ---------------------------------------------------------- 816

Qy   2119 actgcaacttcttattgattaatcaggcaggaggaaggaaaccttggcacagtatcaact 2178
                                    ||||||||||||||||||||||||||||||||||
Db    817 ----------------------caggcaggaggaaggaaaccttggcacagtatcaact 853

Qy   2179 ttgtatgtgcacatgatggatttacactggctgatttggtaacatataataagaagtaca 2238
          |||||||||||||||||||||||||||||| |||||||||||||||||||| |||||||
Db    854 ttgtatgtgcacatgatggatttacactgggtgatttggtaacatataataacaagtaca 913

Qy   2239 atttaccaaatggggagaacaacagagatggagaaaatcacaatcttagctggaattgtg 2298
          |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db    914 atttaccaaatggggagaacaatagagatggagaaaatcacaatcttagctggaattgtg 973

Qy   2299 gggaggaaggagaattcgcaagattgtctgtcaaaagattgaggaagaggcagatgcgca 2358
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    974 gggaggaaggagaattcgcaagattgtctgtcaaaagattgaggaagaggcagatgcgca 1033

Qy   2359 atttctttgtttgtctcatggtttctcaaggagttccaatgttctacatgggtgatgaat 2418
          |||||||||||||||||||||||||||||||||||||||| |||||||| |||||||
Db   1034 atttctttgtttgtctcatggtttctcaaggagttccaatgttttacatgggcgatgaat 1093

Qy   2419 atggccacacaaaaggggcaacaacaatacatactgccatgattcttatgtcaattatt 2478
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1094 atggccacacaaaaggggcaacaacaatacatactgccatgattcttatgtcaattatt 1153

Qy   2479 ttcgctgggataaaaaagaacaatactctgagttgcaccgattctgctgcctcatgacca 2538
```

```
Db   1154 ttcgctgggataaaaaagaacaatactctgacttgcacagattctgctgcctcatgacca 1213

Qy   2539 aattccgcaaggagtgcgagggtcttggccttgaggactttccaacggccaaacggctgc 2598
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db   1214 aattccgcaaggagtgcgagggtcttggccttgaggactttccaacggccgaacggctgc 1273

Qy   2599 agtggcatggtcatcagcctgggaagcctgattggtctgagaatagccgattcgttgcct 2658
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1274 agtggcatggtcatcagcctgggaagcctgattggtctgagaatagccgattcgttgcct 1333

Qy   2659 tttccatgaaagatgaaagacagggcgagatctatgtggccttcaacaccagccacttac 2718
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1334 tttccatgaaagatgaaagacagggcgagatctatgtggccttcaacaccagccacttac 1393

Qy   2719 cggccgttgttgagctcccagagcgcgcagggcgccggtgggaaccggtggtggacacag 2778
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1394 cggccgttgttgagctcccagagcgcgcagggcgccggtgggaaccggtggtggacacag 1453

Qy   2779 gcaagccagcaccatacgacttcctcaccgacgacttacctgatcgcgctctcaccatac 2838
          |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1454 gcaagccagcaccatatgacttcctcaccgacgacttacctgatcgcgctctcaccatac 1513

Qy   2839 accagttctcgcatttcctctactccaacctctaccccatgctcagctactcatcggtca 2898
          |||||||||| |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db   1514 accagttctctcatttcctcaactccaacctctaccccatgctcagctactcatcggtca 1573

Qy   2899 tcctagtattgcgccctgatgtttgagagaccaatatatacagtaaataatatgtctata 2958
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db   1574 tcctagtattgcgccctgatgtttgagagacaaatatatacagtaaataatatgtctata 1633

Qy   2959 tgta 2962
          ||||
Db   1634 tgta 1637
```

```
RESULT   3
AAX76375
ID   AAX76375 standard; cDNA; 2712 BP.
XX
AC   AAX76375;
XX
DT   04-AUG-1999  (first entry)
XX
DE   Zea mays SU1 starch debranching enzyme encoding cDNA clone.
XX
KW   Zea mays; SU1; starch debranching enzyme; maize; endosperm; corn;
KW   bacterial isoamylase; alpha-amylase; starch hydrolytic enzyme;
KW   sugary 1; ss.
XX
OS   Zea mays.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..2457
FT                   /*tag=  a
XX
PN   US5912413-A.
XX
PD   15-JUN-1999.
XX
PF   24-MAR-1995;   95US-0410784.
XX
PR   24-MAR-1995;   95US-0410784.
XX
PA   (IOWA ) UNIV IOWA STATE RES FOUND INC.
XX
PI   James MG,  Myers AM;
XX
DR   WPI; 1999-357234/30.
DR   P-PSDB; AAY17523.
XX
PT   SU1 starch debranching enzyme
XX
PS   Claim 3; Fig 1; 39pp; English.
XX
CC   The present sequence encodes a starch debranching enzyme active in
CC   maize endosperm (Zea mays), designated SU1. SU1 protein is useful as
CC   a replacement for the bacterial and fungal enzymes currently used in
CC   the starch processing industry.
XX
SQ   Sequence 2712 BP; 636 A; 646 C; 745 G; 685 T; 0 other;
```

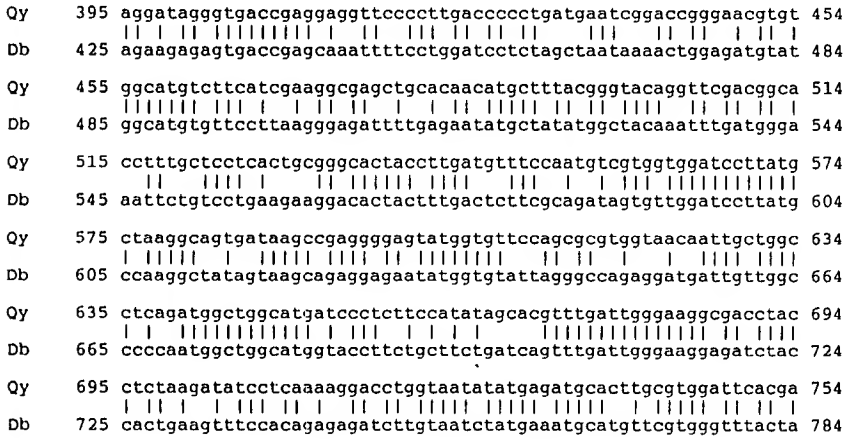
```
  Query Match             33.4%;  Score 1000;  DB 20;  Length 2712;
  Best Local Similarity   65.7%;  Pred. No. 5.4e-206;
  Matches 1883;  Conservative    0;  Mismatches  350;  Indels  633;  Gaps    4;

Qy     79 cggttgtcgaggcggcgacgaaggtagaggacgaggggggaggaggacgagccggtggcgg 138
          | || | |||||||| | | || |||||||| | || || ||||||  |||||| |
Db    233 ccgtggccgaggcggtgcaggcggaggaggacgacgacgacgacgaggaggtggccg 292

Qy    139 aggacaggtacgcgctcggcggcgcgtgcagggtgctcgccggaatgcccgcgccgctgg 198
          |||| |||| |||||| |||||||||||| |||||||||| ||||||||||||||||| |
Db    293 aggagaggttcgcgctgggcggcgcgtgccgggtgctcgcgggaatgcccgcgccgctcg 352

Qy    199 gcgccaccgcgctcgccggcggggtcaatttcgccgtctattccggcggagccaccgccg 258
          ||||||||||||||  |||||| ||||| |||||||||||| ||| |||| ||| ||||||
Db    353 gcgccaccgcgctccgcggcggtgtcaacttcgccgtctactccagcggtgcctccgccg 412

Qy    259 cggcgctctgcctcttcacgccagaagatctcaaggcggtggggttgcctcccgagtaga 318
          || ||||  |||||||| | || |  || ||||||||
Db    413 cgtcgctgagcctcttcgctcccggcgacctcaaggc---------------------- 449

Qy    319 gttcatcagctttgcgtgcgccgcgcgccccttttttgggcctgcaatttaagttttgta 378

Db    450 ---------------------------------------------------------- 449

Qy    379 ctggggcaaatgctgcaggatagggtgaccgaggaggttccccttgaccccctgatgaat 438
                          |||||||||||||||||||| ||||| || |||||| | ||
Db    450 -----------------ggatagggtgaccgaggaggtgcccctcgatcccctgctcaac 492

Qy    439 cggaccgggaacgtgtggcatgtcttcatcgaagg---cgagctgcacaacatgctttac 495
          || || || |||||||||||| || |||||| | ||   ||||||||||  |||||| | |
Db    493 cgaacgggaaacgtgtggcacgtgttcatccacggggacgagctgcacggcatgctctgc 552

Qy    496 gggtacaggttcgacggcacctttgctcctcactgcgggcactaccttgatgtttccaat 555
          || |||||||||| |||   || || ||| |  |||| || |||   |||||  |||||
Db    553 ggatacaggttcgatggcgtgttcgcccctgagcgcggacagtactacgatgtgtccaac 612

Qy    556 gtcgtggtggatccttatgctaaggcagtgataagccgaggggagtatggtgttccagcg 615
          || |||||||||| || |||||||||||| |||||||||| || |||||||| || |||
Db    613 gttgtggtggatccatacgctaaggcagtggtaagccgaggtgaatatggtgtgcctgcg 672

Qy    616 cgtggtaacaattgctggcctcagatggctggcatgatccctcttccatatagcacgttt 675
          | ||||   | ||| |||||||| |||||||| |||||||||||||| ||||  | ||||
Db    673 cctggtggtagttgttggcctcaaatggctggtatgatccctcttccctataataagttt 732

Qy    676 gattgggaaggcgacctacctctaagatatcctcaaaaggacctggtaatatatgagatg 735
          |||||| |||| |||||||| ||  | || | ||| |||||||| || ||||||||  |||
Db    733 gattggcaaggtgacctacccctgggtaccatcagaaggaccttgtcatatatgaaatg 792

Qy    736 cacttgcgtggattcacgaagcatgattcaagcaatgtagaacatccgggtactttcatt 795
          || |||||||||||||| |||||  | ||||||||   | |||| || || |||| ||||
Db    793 catttgcgtggattcacaaagcacaactcaagcaagacaaaacacccaggaacttacatt 852

Qy    796 ggagctgtgtcgaagcttgactatttgaaggagcttggagttaattgtattgaattaatg 855
          || |||||||| ||||||||| || | ||||| |||||||| || ||||| ||  |||||
Db    853 ggtgctgtgtcaaagcttgaccatctaaaggaacttggagtgaactgtatagagctaatg 912

Qy    856 ccctgccatgagttcaacgagctggagtactcaacctcttcttccaagatgaacttttgg 915
          |||||||||||||||| ||||| |||||||  | ||| |||| ||||||||||  |||
Db    913 ccctgccatgagttcaatgagctagagtacttcagctcctcttcgaagatgaacttctgg 972

Qy    916 ggatattctaccataaacttcttttcaccaatgacaagatacacatcaggcgggataaaa 975
          |||||||| || ||||| || || ||||||||| |||||||  | ||| | || |||| |
Db    973 ggatattccacaataaatttttttctcaccaatggcaagatattcttcaagtggcataaga 1032
```

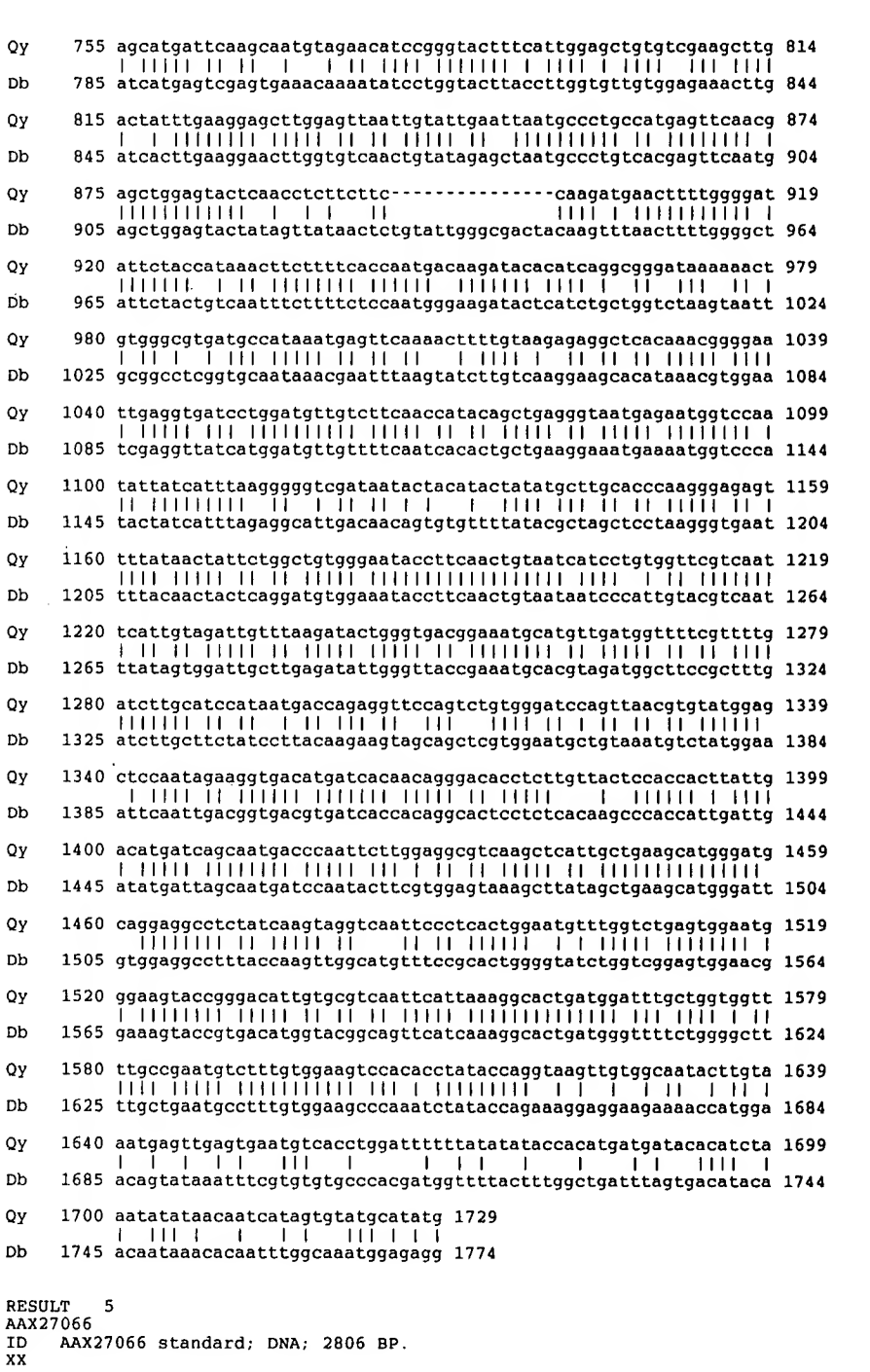
```
Qy    976 aactgtgggcgtgatgccataaatgagttcaaaacttttgtaagagaggctcacaaacgg 1035
Db   1033 gactctggatgtggtgccataaatgaatttaaagcttttgtaagggaggcccacaaacgg 1092

Qy   1036 ggaattgaggtgatcctggatgttgtcttcaaccatacagctgagggtaatgagaatggt 1095
Db   1093 ggaattgaggtgatcatggatgttgtcttcaatcatacagctgaaggtaatgagaaaggc 1152

Qy   1096 ccaatattatcatttaaggggtcgataatactacatactatatgcttgcacccaaggga 1155
Db   1153 ccaatattatcctttaggggatagataatagtacatactacatgcttgcacctaaggga 1212

Qy   1156 gagttttataactattctggctgtgggaataccttcaactgtaatcatcctgtggttcgt 1215
Db   1213 gagttttataattattctggttgtggaaataccttcaattgtaatcatcctgtagtccgt 1272

Qy   1216 caattcattgtagattgtttaagatactgggtgacggaaatgcatgttgatggttttcgt 1275
Db   1273 gaatttatagtggattgcttgagatactgggtaacagaaatgcatgttgatggttttcgt 1332

Qy   1276 tttgatcttgcatccataatgaccagaggttccagtctgtgggatccagttaacgtgtat 1335
Db   1333 tttgaccttgcatctatactgaccagaggatgcagtctatgggatccagttaatgtgtat 1392

Qy   1336 ggagctccaatagaaggtgacatgatcacaacagggacacctcttgttactccaccactt 1395
Db   1393 ggaagtccaatggaaggtgacatgattacgacagggacacctcttgttgccccaccactt 1452

Qy   1396 attgacatgatcagcaatgacccaattcttggaggcgtcaagctcattgctgaagcatgg 1455
Db   1453 attgacatgattagcaatgacccaattcttggaaatgtcaagctcattgctgaagcatgg 1512

Qy   1456 gatgcaggaggcctctatcaagtaggtcaattccctcactggaatgtttggtctgagtgg 1515
Db   1513 gatgcaggaggtctctatcaagaaggtcagtttcctcactggaacgtttggtcagagtgg 1572

Qy   1516 aatgggaagtaccgggacattgtgcgtcaattcattaaaggcactgatggatttgctggt 1575
Db   1573 aatggaaagtatcgcgataccgtgcgtcagttcatcaaaggcacagatggatttgctggt 1632

Qy   1576 ggttttgccgaatgtctttgtggaagtccacacctataccaggtaagttgtggcaatact 1635
Db   1633 gcttttgctgaatgcctatgtggaagtccacagttatac---------------------- 1671

Qy   1636 tgtaaatgagttgagtgaatgtcacctggatttttatatataccacatgatgatacaca 1695
Db   1672 ------------------------------------------------------------ 1671

Qy   1696 tctaaatatataacaatcatagtgtatgcatatgcatttggctaagaagtattagtgtat 1755
Db   1672 ------------------------------------------------------------ 1671

Qy   1756 acactagtgctatatataggttttaacacccaacttgccaatgaaggaacatagggcttt 1815
Db   1672 ------------------------------------------------------------ 1671

Qy   1816 ctagttatcttatttatttgtccggtgaataatccactgaaaaattccagccatgtcatt 1875
Db   1672 ------------------------------------------------------------ 1671

Qy   1876 ttttaggggggagaagaaactatattgatttgccccccctaaaagaagccatctcagaat 1935
Db   1672 ------------------------------------------------------------ 1671

Qy   1936 tcataggtaagttgctttctgtaaagaaaggaaaacgacttcatactttctatcggtgc 1995
Db   1672 ------------------------------------------------------------ 1671

Qy   1996 taacttagctcgatgtatatttgtaagatgaatgccaaatttaatttgtcggataatttg 2055
Db   1672 ------------------------------------------------------------ 1671

Qy   2056 atctgttattcacaaatttctatttggtttctctagaaatcaaaccagtaacttgttatt 2115
Db   1672 ------------------------------------------------------------ 1671

Qy   2116 ggcactgcaacttcttattgattaatcaggcaggaggaaggaaaccttggcacagtatca 2175
Db   1672 -------------------------caggcaggggggaggaagccttggcacagtatcg 1705

Qy   2176 actttgtatgtgcacatgatggatttacactggctgatttggtaacatataataagaagt 2235
Db   1706 gctttgtatgtgcacacgatggatttacactggctgatttggtcacatacaatagcaagt 1765

Qy   2236 acaatttaccaaatggggagaacaacagagatggagaaaatcacaatcttagctggaatt 2295
Db   1766 acaacttgtcaaatggtgaggacttcagagatggggaaaatcataatcttagctggaatt 1825

Qy   2296 gtggggaggaaggagaattcgcaagattgtctgtcaaaagattgaggaagaggcagatgc 2355
Db   1826 gtggggaggaaggagaatttgcaagtctgtcagtccgaagattaaggaagaggcaaatgc 1885

Qy   2356 gcaatttctttgtttgtctcatggtttctcaaggagttccaatgttctacatgggtgatg 2415
Db   1886 gcaatttctttgtttgtcttatggtttctcagggagttccaatgttctacatgggcgatg 1945

Qy   2416 aatatggccacacaaaaggggcaacaacaatacatactgccatgattcttatgtcaatt 2475
Db   1946 aatatggtcacacaaagggagggaacaacaatacgtactgccatgaccattatgtcaatt 2005

Qy   2476 attttcgctgggataaaa---aagaacaatactctgagttgcaccgattctgctgcctca 2532
Db   2006 atttccgttgggataagaaggaagaacaatcctctgatttgtacagattctgccgtctca 2065

Qy   2533 tgaccaaattccgcaaggagtgcgagggtcttggccttgaggactttccaacggccaaac 2592
Db   2066 tgaccgaattccgcaaagaatgtgaatctcttggccttgaggacttcccgacttcagaac 2125

Qy   2593 ggctgcagtggcatggtcatcagcctgggaagcctgattggtctgagaatagccgattcg 2652
Db   2126 ggttgaaatggcacggtcatcagcccgggaagcctgactggtcagaggcaagccgattcg 2185

Qy   2653 ttgccttttccatgaaagatgaaagacagggcgagatctatgtggccttcaacaccagcc 2712
Db   2186 ttgccttcaccatgaaggacgaaaccaaaggcgagatctacgtggccttcaacaccagtc 2245

Qy   2713 acttaccggccgttgttgagctcccagagcgcgcagggcgccggtgggaaccggtggtgg 2772
Db   2246 accttccggtggttgttgggcttccagagcgctctgggttccgatgggagccggtggtgg 2305

Qy   2773 acacaggcaagccagcaccatacgacttcctcaccgacgacttacctgatcgcgctctca 2832
Db   2306 acaccggcaaggaggcaccatatgacttcctcaccgatggcctgccagatcgtgctgtca 2365

Qy   2833 ccatacaccagttctcgcatttcctctactccaacctctaccccatgctcagctactcat 2892
Db   2366 ccgtctaccagttctctcatttcctcaactccaatctctatcctatgctcagctactcct 2425

Qy   2893 cggtcatcctagtattgcgccctgatgtttgagagaccaatatata 2938
Db   2426 ccatcatccttgtattgcgccctgatgtctgaaagaagcagataca 2471
```

RESULT 4
AAX27061
ID AAX27061 standard; DNA; 2706 BP.
XX
AC AAX27061;
XX
DT 21-MAY-1999 (first entry)
XX
DE S. tuberosum isoamylase clone 15 coding sequence.
XX
KW Isoamylase clone; potato; branched polysaccharide synthesis; amylopectin;
KW starch modification; debranching activity; ss.

XX
OS Solanum tuberosum.
XX
PN WO9906575-A1.
XX
PD 11-FEB-1999.
XX
PF 30-JUL-1998; 98WO-GB02280.
XX
PR 31-JUL-1997; 97GB-0016185.
XX
PA (PLAN-) PLANT BIOSCIENCE LTD.
XX
PI Bustos Guillen R, Edwards EA, Martin CR, Smith AM;
XX
DR WPI; 1999-153803/13.
DR P-PSDB; AAY00869.
XX
PT New nucleic acid from potato encoding starch debranching enzyme with
PT isoamylase activity - and related vectors, transformed cells and
PT plants, proteins and antibodies, used to generate starch with
PT altered properties, for use in foods, as thickeners etc.
XX
PS Claim 3; Fig 2; 78pp; English.
XX
CC This sequence encodes a Solanum tuberosum isoamylase of the invention.
CC Fragments of the isoamylase coding sequence are used to identify and
CC clone isoamylases from other plant species, by standard hybridisation
CC or amplification methods. Expression of the isoamylase DNA in host cells
CC is used to produce the isoamylase which is used: (i) to raise antibodies
CC (Ab) (for identification, isolation and localisation of isoamylases); and
CC (ii) for synthesis of branched polysaccharides. The DNA, proteins
CC and Ab (or related peptides) are used to alter the quality and
CC quantity of polysaccharides in a host cells, specifically to alter
CC branching in amylopectin. The modified starches produced are useful in
CC human or animal foods; as biodegradable plastic; as food or paint
CC thickener; in starch-coated films, paper and textiles; in mining
CC explosives; in pharmaceuticals and glues. Manipulation of debranching
CC activity in a plant allows control of starch properties, e.g. increased
CC gel strength; formation of paste rather than gel, changed physical
CC characteristics etc.
XX
SQ Sequence 2706 BP; 797 A; 492 C; 624 G; 792 T; 1 other;

Query Match 22.5%; Score 673.2; DB 20; Length 2706;
Best Local Similarity 69.4%; Pred. No. 1.8e-135;
Matches 937; Conservative 0; Mismatches 398; Indels 15; Gaps 1;





RESULT 5
AAX27066
ID AAX27066 standard; DNA; 2806 BP.
XX

```
AC   AAX27066;
XX
DT   21-MAY-1999  (first entry)
XX
DE   Original S. tuberosum isoamylase clone 15 coding sequence.
XX
KW   Isoamylase clone; potato; branched polysaccharide synthesis; amylopectin;
KW   starch modification; debranching activity; ss.
XX
OS   Solanum tuberosum.
XX
PN   WO9906575-A1.
XX
PD   11-FEB-1999.
XX
PF   30-JUL-1998;   98WO-GB02280.
XX
PR   31-JUL-1997;   97GB-0016185.
XX
PA   (PLAN-) PLANT BIOSCIENCE LTD.
XX
PI   Bustos Guillen R,  Edwards EA,  Martin CR,  Smith AM;
XX
DR   WPI; 1999-153803/13.
DR   P-PSDB; AAY00872.
XX
PT   New nucleic acid from potato encoding starch debranching enzyme with
PT   isoamylase activity - and related vectors, transformed cells and
PT   plants, proteins and antibodies, used to generate starch with
PT   altered properties, for use in foods, as thickeners etc.
XX
PS   Disclosure; Fig 11; 78pp; English.
XX
CC   This sequence represents the original isolated clone encoding a
CC   Solanum tuberosum isoamylase of the invention.
CC   Fragments of the isoamylase coding sequence are used to identify and
CC   clone isoamylases from other plant species, by standard hybridisation
CC   or amplification methods. Expression of the isoamylase DNA in host cells
CC   is used to produce the isoamylase which is used: (i) to raise antibodies
CC   (Ab) (for identification, isolation and localisation of isoamylases); and
CC   (ii) for synthesis of branched polysaccharides. The DNA, proteins
CC   and Ab (or related peptides) are used to alter the quality and
CC   quantity of polysaccharides in a host cells, specifically to alter
CC   branching in amylopectin. The modified starches produced are useful in
CC   human or animal foods; as biodegradable plastic; as food or paint
CC   thickener; in starch-coated films, paper and textiles; in mining
CC   explosives; in pharmaceuticals and glues. Manipulation of debranching
CC   activity in a plant allows control of starch properties, e.g. increased
CC   gel strength; formation of paste rather than gel, changed physical
CC   characteristics etc.
XX
SQ   Sequence 2806 BP; 818 A; 523 C; 654 G; 807 T; 4 other;

  Query Match             20.5%;  Score 613.2;  DB 20;  Length 2806;
  Best Local Similarity   68.0%;  Pred. No. 1.6e-122;
  Matches  922;  Conservative    0;  Mismatches  410;  Indels   24;  Gaps    4;

Qy    395 aggatagggtgaccgaggaggttccccttgacccccgatgaatcggaccgggaacgtgt 454
          || | || |||||||||| |  ||  ||| || || ||    |||   || ||  | || |
Db    488 agaagagagtgaccgagcaaattttcctggatcctctagctaataaaactggagatgtat 547

Qy    455 ggcatgtcttcatcgaaggcgagctgcacaacatgctttacgggtacaggttcgacggca 514
          ||||||| ||| |  | || ||  |  | || ||||| || || |||| || || | |
Db    548 ggcatgtgttccttaagggagattttgagaatatgctatatggctacaaatttgatggga 607

Qy    515 cctttgctcctcactgcgggcactaccttgatgtttccaatgtcgtggtggatccttatg 574
            ||   |||| |    || |||||| ||||   |||  |  | ||| ||||||||||||
Db    608 aattctgtcctgaagaaggacactactttgactcttcgcagatagtgttggatccttatg 667

Qy    575 ctaaggcagtgataagccgaggggagtatggtgttccagcgcgtggtaacaattgctggc 634
          | |||||  |  ||||| |||| || |||||||||   || ||  |   |  |||| ||||
Db    668 ccaaggctatagtaagcagaggagaatatggtgtattagggccagaggatgattgttggc 727

Qy    635 ctcagatggctggcatgatccctct---tccatatagcacgtttgattgggaaggcgacc 691
          | |  ||||||||||||| | ||  |       |  | ||  ||  |||||||||| || |
Db    728 ccccaatggctggcatggtacccttctgcttctggatcagtttgtattgggaaggagatc 787

Qy    692 tacctctaagatatcctca---aaaggacctggtaatatatgagatgcacttgcgtggat 748
          |||| ||   |  | | ||   | ||  | || | ||  |||| ||||  | |||||
Db    788 taccactggaagtttcccacagagagatcttgttnatcnatgaaatgcatgttcgtgggt 847

Qy    749 tcacgaagcatgattcaagcaatgtagaacatccgggtactttcattggagctgtgtcga 808
          | || |  ||||| || ||  |  | || |||| |||||| | |||| | ||||  ||
Db    848 ttactatccatgagtcgagtgaaacaaaatatcctggtacttaccttggtgttgtggaga 907

Qy    809 agcttgactatttgaaggagcttggagttaattgtattgaattaatgccctgccatgagt 868
          | |||||  | |||||||| ||||| || || ||||| ||  |||||||||| || ||||
Db    908 aacttgatcacttgaaggaacttggtgtcaactgtatagagctaatgccctgtcacgagt 967

Qy    869 tcaacgagctggagtactcaacctcttcttc----------------caagatgaacttttt 913
          |||| ||||||||||||||  |  | |  ||               |||| | |||||||
Db    968 tcaatgagctggagtactatagttataactctgtattgggcgactacaagtttaacttttt 1027

Qy    914 ggggatattctaccataaacttcttttcaccaatgacaagatacacatcaggcgggataa 973
          |||| ||||||||  | || ||||||||| ||||||  |||||||| |||| |  ||  |||
Db   1028 ggggctattctactgtcaatttctttttctccaatgggaagatactcatctgctggtctaa 1087

Qy    974 aaaactgtgggcgtgatgccataaatgagttcaaaacttttgtaagagaggctcacaaac 1033
            || || || |  | ||| ||||| || || ||   | |||| |  || || || ||||
Db   1088 gtaattgcggcctcggtgcaataaacgaatttaagtatcttgtcaaggaagcacataaac 1147

Qy   1034 ggggaattgaggtgatcctggatgttgtcttcaaccatacagctgagggtaatgagaatg 1093
          | ||||| ||||| ||| |||||||||| ||||| || || ||||| || ||||| ||||
Db   1148 gtggaatcgaggttatcatggatgttgttttcaatcacactgctgaaggaaatgaaaatg 1207

Qy   1094 gtccaatattatcatttaagggggtcgataatactacatactatatgcttgcacccaagg 1153
          |||| ||| |||||||||  ||  | || || | |   |  |||| ||| || || ||||
Db   1208 gtcccatactatcatttagaggcattgacaacagtgtgttttatacgctagctcctaagg 1267

Qy   1154 gagagttttataactattctggctgtgggaataccttcaactgtaatcatcctgtggttc 1213
          | || ||||| ||||| || || ||||| ||||||||||||||||||| ||||  | || |
Db   1268 gtgaattttacaactactcaggatgtggaaataccttcaactgtaataatcccattgtac 1327

Qy   1214 gtcaattcattgtagattgtttaagatactgggtgacggaaatgcatgttgatggttttc 1273
          ||||||   || |||       | ||||| ||||| || |||||||| ||  |||| || |
Db   1328 gtcaat---ttatagtgatgctgagatattgggttaccgaaatgcacgtacatggcttcc 1384

Qy   1274 gttttgatcttgcatccataatgaccagaggttccagtctgtgggatccagttaacgtgt 1333
          | |||||||||| || ||  | || ||| ||  |||   |||| || | || || || |
Db   1385 gctttgatcttgcttctatccttacaagaagtagcagctcgtggaatgctgtaaatgtct 1444

Qy   1334 atggagctccaatagaaggtgacatgatcacaacagggacacctcttgttactccaccac 1393
          |||||  | |||| || |||||| ||||||| ||||| || |||||    |  ||||||
Db   1445 atggaaattcaattgacggtgacgtgatcaccacaggcactcctctcacaagcccaccat 1504

Qy   1394 ttattgacatgatcagcaatgacccaattcttggaggcgtcaagctcattgctgaagcat 1453
          | ||||| ||||| |||||||| ||||| ||| | || || ||||| || ||||||||||
Db   1505 tgattgatatgattagcaatgatccaatacttcgtggagtaaagcttatagctgaagcat 1564

Qy   1454 gggatgcaggaggcctctatcaagtaggtcaattccctcactggaatgtttggtctgagt 1513
          |||||   |||||||| || ||||| ||    || || ||||||  | | ||||| ||||
Db   1565 gggattgtggaggcctttaccaagttggcatgtttccgcactggggtatctggtcggagt 1624

Qy   1514 ggaatgggaagtaccgggacattgtgcgtcaattcattaaaggcactgatggatttgctg 1573
          |||| || |||||||| ||||| || || || ||||| |||||||||||||| ||| |||
Db   1625 ggaacggaaagtaccgtgacatggtacggcagttcatcaaaggcactgatgggttttctg 1684

Qy   1574 gtggttttgccgaatgtctttgtggaagtccacacctataccaggtaagttgtggcaata 1633
          | | |||||| ||||| |||||||||| ||| | |||||||||  | |  |  | ||
Db   1685 gggcttttgctgaatgcctttgtggaagcccaaatctataccagaaaggaggaagaaaac 1744

Qy   1634 cttgtaaatgagttgagtgaatgtcacctggatttttatatataccacatgatgataca 1693
          | || ||  |  | | |  |||  |     |  | |  |   |  |  | |  |
Db   1745 catggaacagtataaatttcgtgtgtgcccacgatggttttactttggctgatttagtga 1804
```

```
Qy   1694 catctaaatatataacaatcatagtgtatgcatatg 1729
          ||| ||  ||| |   |   | |   ||| | | |
Db   1805 catacaacaataaacacaatttggcaaatggagagg 1840


RESULT   6
AAT96754
ID   AAT96754 standard; cDNA to mRNA; 2133 BP.
XX
AC   AAT96754;
XX
DT   08-MAY-1998  (first entry)
XX
DE   S. tuberosum debranching enzyme cDNA.
XX
KW   Debranching enzyme; PDE; potato; transgenic plant; starch; industry;
KW   food manufacture; ss.
XX
OS   Solanum tuberosum.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1822
FT                   /*tag= a
FT                   /product= debranching enzyme
FT                   /note= "partial coding sequence"
XX
PN   DE19618125-A1.
XX
PD   13-NOV-1997.
XX
PF   06-MAY-1996;   96DE-1018125.
XX
PR   06-MAY-1996;   96DE-1018125.
XX
PA   (PLAN-) PLANTEC BIOTECHNOLOGIE GMBH.
XX
DR   WPI; 1997-551167/51.
DR   P-PSDB; AAW36602.
XX
PT   DNA encoding potato debranching enzyme - useful for producing
PT   recombinant enzyme or transgenic plants for use in food production
PT   or in industry
XX
PS   Claim 1; Page 14-17; 20pp; German.
XX
CC   This cDNA sequence encodes a novel potato debranching enzyme (PDE).
CC   This gene can be used in the production of transgenic plants especially
CC   starch storing plants e.g. Cereal or potato plants. Other nucleic acid
CC   molecules could be designed which specifically hybridise to the PDE
CC   transcript and inhibit its translation. Vectors containing this gene can
CC   be linked in sense orientation with regulatory elements that ensure its
CC   transcription and translation in prokaryotic and eukaryotic cells. The
CC   starch produced from such transgenic plants could be used in the
CC   manufacture of foods or industrial products.
XX
SQ   Sequence 2133 BP; 650 A; 392 C; 475 G; 616 T; 0 other;


  Query Match             19.2%;  Score 575.4;  DB 18;  Length 2133;
  Best Local Similarity   69.8%;  Pred. No. 2.1e-114;
  Matches  799;  Conservative    0;  Mismatches  331;  Indels   15;  Gaps    1;

Qy    600 gtatggtgttccagcgcgtggtaacaattgctggcctcagatggctggcatgatcccctct 659
          | ||   |  | || ||  |   |  |||| ||||| |  ||||| |||||| | |||
Db      1 gaattcggcacgagggccagaggatgattgttggccccccaatggcaggcatggtaccttc 60

Qy    660 tccatatagcacgtttgattgggaaggcgacctacctctaagatatcctcaaaaggacct 719
          | | | |     |||||||||||||||| || |||   || |  | ||| || |  || ||
Db     61 tgcttctgatcagtttgattgggaaggagatctattactgaagtttccacagagagatct 120

Qy    720 ggtaatatatgagatgcacttgcgtggattcacgaagcatgattcaagcaatgtagaaca 779
          |||||  |||||  |||||  | ||||||| || || ||||| || ||  |   | || |
Db    121 tgtaatctatgaaatgcatgttcgtggatttacaaatcatgagtcgagtgaaacaaaata 180

Qy    780 tccgggtactttcattggagctgtgtcgaagcttgactatttgaaggagcttggagttaa 839
          ||| |||||||| | |||| | ||||  ||| |||||  |. |||||||| ||||| || ||
Db    181 tcctggtacttaccttggtgttgtggagaaacttgatcacttgaaggaacttggtgtcaa 240

Qy    840 ttgtattgaattaatgccctgccatgagttcaacgagctggagtactcaacctcttcttc 899
          |||||  ||  |||||||||| || |||||||| ||||||||||||||  |  | |  ||
Db    241 ctgtatagagctaatgccctgtcacgagttcaatgagctggagtactatagttataactc 300

Qy    900 ---------------caagatgaacttttggggatattctaccataaacttcttttcacc 944
                          |||| | |||||||||| ||||||||  | || |||||||| ||
Db    301 tgtattgggcgactacaagtttaacttttggggctattctactgtcaatttcttttctcc 360

Qy    945 aatgacaagatacacatcaggcgggataaaaaactgtgggcgtgatgccataaatgagtt 1004
          ||||  |||||| | || |  ||  |||  || || || |  | ||| ||||| || ||
Db    361 aatgggaagatactcgtctgctggtctaagtaattgcggcctcggtgcaataaacgaatt 420

Qy   1005 caaaacttttgtaagagaggctcacaaacggggaattgaggtgatcctggatgttgtctt 1064
           ||   | |||| |  || || || ||||| ||||| ||||| ||| ||||||||| ||
Db    421 taagtatcttgtcaaggaagcacataaacgtggaatcgaggttatcatggatgttgtttt 480

Qy   1065 caaccatacagctgagggtaatgagaatggtccaatattatcatttaaggggtcgataa 1124
          ||| || || ||||| || ||||| |||||||| ||| |||||||||  ||  | || ||
Db    481 caatcacactgctgaaggaaatgaaaatggtcccatactatcatttagaggcattgacaa 540

Qy   1125 tactacatactatatgcttgcacccaagggagagttttataactattctggctgtgggaa 1184
           | |   |  |||| ||| || || ||||| || ||||| ||||| || || ||||| ||
Db    541 cagtgtgttttatacgctagctcctaagggtgaattttacaactactcaggatgtggaaa 600

Qy   1185 taccttcaactgtaatcatcctgtggttcgtcaattcattgtagattgtttaagatactg 1244
          ||||||||||||||||| ||||  | || |||||||| || || ||||| || ||||| ||
Db    601 taccttcaactgtaataatcccattgtacgtcaatttatagtggattgcttgagatattg 660

Qy   1245 ggtgacggaaatgcatgttgatggttttcgttttgatcttgcatccataatgaccagagg 1304
          ||| || |||||||| || ||||| || || |||||||||| || ||  | || ||| |
Db    661 ggttaccgaaatgcacgtagatggcttccgctttgatcttgcttctatccttacaagaag 720

Qy   1305 ttccagtctgtgggatccagttaacgtgtatggagctccaatagaaggtgacatgatcac 1364
          |  |||   |||| || | || || || ||||||  | |||| || ||||||||||||||
Db    721 tagcagctcgtggaatgctgtaaatgtctatggaaattcaattgacggtgacatgatcac 780

Qy   1365 aacagggacacctcttgttactccaccacttattgacatgatcagcaatgacccaattct 1424
           ||||| || |||||    |  |||||| | ||||| ||||| |||||||| ||||| ||
Db    781 cacaggcactcctctcacaagcccaccattgattgatatgattagcaatgatccaatact 840

Qy   1425 tggaggcgtcaagctcattgctgaagcatgggatgcaggaggcctctatcaagtaggtca 1484
          | | || || ||||| || ||||||||||||||||    |||||||| || ||||| ||
Db    841 tagtggagtaaagcttatagctgaagcatgggattgtggaggcctttaccaagttggcat 900

Qy   1485 attccctcactggaatgtttggtctgagtggaatgggaagtaccgggacattgtgcgtca 1544
           || || ||||||  | | ||||| |||||||| || |||||||| ||||| || |||||
Db    901 gtttccgcactggggtatctggtcggagtggaacggaaagtaccgtgacatggtacgtca 960

Qy   1545 attcattaaaggcactgatggatttgctggtggttttgccgaatgtctttgtggaagtcc 1604
           ||||| ||||||||||||||| ||| |||| | |||||| ||||| |||||||||| ||
Db    961 gttcatcaaaggcactgatgggttttctggggcttttgctgaatgcctttgtggaagccc 1020

Qy   1605 acacctataccaggtaagttgtggcaatacttgtaaatgagttgagtgaatgtcacctgg 1664
          | | |||||||||  | |  |  | ||  | || ||  |  |  | |  |||  |
Db    1021 aaatctataccagaaaggaggaagaaaaccatggaacagtataaatttcgtgtgtgccca 1080

Qy   1665 atttttatatataccacatgatgatacacatctaaatatataacaatcatagtgtatgc 1724
             |  | |   |    |    | |   |||| ||  ||| |   |   | |  |||
Db    1081 cgatggttttacttggctgatttagtgacatacaacaataaacacaatttggcaaatgg 1140

Qy   1725 atatg 1729
          | | |
Db    1141 agagg 1145


RESULT   7
AAX34655
ID   AAX34655 standard; DNA; 9289 BP.
```

```
XX
AC   AAX34655;
XX
DT   05-JUL-1999  (first entry)
XX
DE   Hexaploid wheat DBE genomic sequence.
XX
KW   Starch biosynthetic pathway; cereal plant; enzyme; SBE; SSS; DBE; GBSS;
KW   starch branching enzyme; starch soluble synthase; debranching enzyme;
KW   endosperm; wheat; barley; granule-bound synthase; glutenin; starch;
KW   grain softness protein I; bacterial isoamylase; glycogen synthase;
KW   wSBE I-D4 gene; ss.
XX
OS   Triticum tauschii.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..9287
FT                   /*tag= a
FT                   /note= "the start codon is not indicated; contains
FT                          146 internal stop codons"
XX
PN   WO9914314-A1.
XX
PD   25-MAR-1999.
XX
PF   11-SEP-1998;   98WO-AU00743.
XX
PR   20-MAR-1998;   98AU-0002509.
PR   12-SEP-1997;   97AU-0009108.
XX
PA   (CSIR ) COMMONWEALTH SCI & IND RES ORG.
PA   (GOOD-) GOODMAN FIELDER LTD.
PA   (LIMA-) GRP LIMAGRAIN PACIFIC PTY LTD.
PA   (AUSU ) UNIV AUSTRALIAN NAT.
XX
PI   Li Z,  Morell M,  Rahman S;
XX
DR   WPI; 1999-229525/19.
DR   P-PSDB; AAY06919.
XX
PT   New isolated cereal plant enzyme genes used for, e.g. expression of
PT   antisense sequences of granule bound synthase
XX
PS   Claim 14; Page 103-115; 171pp; English.
XX
CC   The invention relates to a novel enzyme of starch biosynthetic pathway
CC   in a cereal plant, where the enzyme is selected from starch branching
CC   enzyme (SBE) I , SBE II, starch soluble synthase (SSS) I, and debranching
CC   enzyme (DBE), with the proviso that the enzyme is not SSS I of rice, or
CC   SBE I of rice or maize. The methods and products can be used for
CC   targeting expression specifically to the endosperm of the seeds of cereal
CC   plants such as wheat or barley. They can be used for the expression of
CC   e.g. antisense sequences of granule-bound synthase (GBSS), SBE II, low
CC   mol. wt. glutenin, grain softness protein I, bacterial isoamylase,
CC   bacterial glycogen synthase, and wheat high mol. wt. glutenin Bx17. They
CC   can be used for modifying the characteristics of starch produced by a
CC   plant. The present sequence represents the wheat DBE genomic DNA
CC   sequence.
XX
SQ   Sequence 9289 BP; 2477 A; 1909 C; 2040 G; 2863 T; 0 other;


  Query Match             17.8%;  Score 533.4;  DB 20;  Length 9289;
  Best Local Similarity   95.1%;  Pred. No. 3.9e-105;
  Matches  562;  Conservative    0;  Mismatches   26;  Indels    3;  Gaps    1;

Qy       1 ggtcggggccggcgccgcgcctgcgacggtggcgacccaatgcgacggcggggaaggggg 60
           || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1943 ggccggggccggcgccgcgcctgcgacggtggcgacccaatgcgacggcggggaaggggg 2002

Qy      61 tcggcgaggtgtgcgccgcggttgtcgaggcggcgacgaaggtagaggacgag---gggg 117
           ||||||||||||||||||||||||||||||||||||||||  ||||| |||   |  |
Db    2003 tcggcgaggtgtgcgccgcggttgtcgaggcggcgacgaaggccgaggatgaggacgacg 2062

Qy     118 aggaggacgagccggtggcggaggacaggtacgcgctcggcggcgcgtgcagggtgctcg 177
           | ||||| ||| ||||||||||||||||||||||||||||||||||||||||||||||
Db    2063 acgaggaggaggcggtggcggaggacaggtacgcgctcggcggcgcgtgcagggtgctcg 2122

Qy     178 ccggaatgcccgcgccgctgggcgccaccgcgctcgccggcggggtcaatttcgccgtct 237
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2123 ccggaatgcccgcgccgctgggcgccaccgcgctcgccggcggggtcaatttcgccgtct 2182

Qy     238 attccggcggagccaccgccgcggcgctctgcctcttcacgccagaagatctcaaggcgg 297
           | ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2183 actccggtggagccaccgccgcggcgctctgcctcttcacgccagaagatctcaaggcgg 2242

Qy     298 tggggttgcctcccgagtagagttcatcagctttgcgtgcgccgcgcgccccttttttgg 357
           |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||  |
Db    2243 tggggttgcctcccgagtagagttcatcagctttgcgtgcgccgcgcgccccctttttctg 2302

Qy     358 gcctgcaatttaagttttgtactggggcaaatgctgcaggatagggtgaccgaggaggtt 417
           |||||| |||||||||||||||||| |||||||||||||||||||||| |||||||||
Db    2303 gcctgcgatttaagttttgtactgggggaaatgctgcaggatagggtgacggaggaggtt 2362

Qy     418 ccccttgacccccctgatgaatcggaccgggaacgtgtggcatgtcttcatcgaaggcgag 477
           |||||||||||||||||||||||| |||||||||||||||||||||| |||||||||
Db    2363 tcccttgacccccctgatgaatcggactgggaacgtgtggcatgtcttcattgaaggcgag 2422

Qy     478 ctgcacaacatgctttacgggtacaggttcgacggcacctttgctcctcactgcgggcac 537
           |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2423 ctgcacgacatgctttacgggtacaggttcgacggcacctttgctcctcactgcgggcac 2482

Qy     538 taccttgatgtttccaatgtcgtggtggatccttatgctaaggcagtgata 588
           ||||||||| |||||||||||||||||||||||||||||||    | |||
Db    2483 taccttgatatttccaatgtcgtggtggatccttatgctaaggtgatcata 2533


RESULT   8
AAX33162
ID   AAX33162 standard; DNA; 1663 BP.
XX
AC   AAX33162;
XX
DT   06-JUL-1999  (first entry)
XX
DE   Potato isoamylase type DBE genomic clone SEQ ID NO:14.
XX
KW   Potato; isoamylase debranching enzyme; DBE; starch; ss.
XX
OS   Solanum tuberosum.
XX
PN   WO9912950-A2.
XX
PD   18-MAR-1999.
XX
PF   04-SEP-1998;   98WO-GB02665.
XX
PR   06-SEP-1997;   97GB-0018863.
XX
PA   (NATT ) NAT STARCH & CHEM INVESTMENT HOLDING COR.
XX
PI   Jobling SA,  Schwall GP,  Westcott RJ;
XX
DR   WPI; 1999-229220/19.
DR   P-PSDB; AAY04382.
XX
PT   New isolated potato isoamylase-type debranching enzyme gene
XX
PS   Claim 7; Fig 8; 72pp; English.
XX
CC   The present invention describes a novel nucleic acid sequence which is
CC   obtainable from potato plants and carries at least a portion of an
CC   isoamylase-type debranching enzyme (DBE) gene. Constructs containing the
CC   nucleic acid sequence can be used to alter the starch properties of
CC   plants such as potato, sweet potato, maize, wheat, barley, oat, cassava,
CC   pea or rice. The starch can have increased branching and/or shorter
```

```
CC  chain length, reduced peak viscosity, higher setback viscosity or
CC  increased viscosity onset temperature. By using an antisense sequence
CC  with greater homology to the native gene, greater inhibition can be
CC  achieved.
XX
SQ  Sequence 1663 BP; 503 A; 312 C; 381 G; 467 T; 0 other;


  Query Match             13.5%;  Score 405.8;  DB 20;  Length 1663;
  Best Local Similarity   70.7%;  Pred. No. 7.5e-78;
  Matches  555;  Conservative    0;  Mismatches  227;  Indels    3;  Gaps    1;

Qy   2140 atcaggcaggaggaaggaaaccttggcacagtatcaactttgtatgtgcacatgatggat 2199
           | |||  |||||||||| ||||| ||| ||||||| || || || ||||| || |||||
Db    577 accagaaaggaggaagaaaaccatggaacagtataaatttcgtgtgtgcccacgatggtt 636

Qy   2200 ttacactggctgatttggtaacatataataagaagtacaatttaccaaatggggagaaca 2259
           ||||  |||||||||| || ||||| || || ||  ||||||| |||||| ||| |||
Db    637 ttactttggctgatttagtgacatacaacaataaacacaatttggcaaatggagaggaca 696

Qy   2260 acagagatggagaaaatcacaatcttagctggaattgtggggaggaaggagaattcgcaa 2319
           ||| ||| || || |||||||||  ||| ||||||||||| |||||||||||||| ||||
Db    697 acaaagacggggagaatcacaataatagttggaattgtggtgaggaaggagaatttgcaa 756

Qy   2320 gattgtctgtcaaaagattgaggaagaggcagatgcgcaatttctttgtttgtctcatgg 2379
             | | | | || | |||||||||| || || ||||| || |||||  | || || ||||
Db    757 ctatctttatgaagaaattgaggaaaagacaaatgcggaacttcttcctctgccttatgg 816

Qy   2380 tttctcaaggagttccaatgttctacatgggtgatgaatatggccacacaaaaggggca 2439
           |||| ||||| ||||| ||| | || ||||| |||||||||| ||||| || || || |
Db    817 tttcccaaggtgttcccatgatatatatgggcgatgaatatggtcacactaagggaggaa 876

Qy   2440 acaacaatacatactgccatgattcttatgtcaattatttttcgctgggataaaaaagaac 2499
           ||||||| || || |||||||||  |||| | ||||| || || |||||||| || ||
Db    877 acaacaacacgtattgccatgataattatattaattacttccgttgggataagaaggatg 936

Qy   2500 aa---tactctgagttgcaccgattctgctgcctcatgaccaaattccgcaaggagtgcg 2556
           ||   |  ||||| ||     |||| ||| ||||||||||||||||||| | || || |
Db    937 aatcttcatctgatttttgagattttgcggcctcatgaccaaattccgccatgaatgtg 996

Qy   2557 agggtcttggccttgaggactttccaacggccaaacggctgcagtggcatggtcatcagc 2616
           |    || ||  | || |  || || || ||  || ||||||| ||||||||||    |
Db    997 aatcactgggattagatggtttccctacagcagaaaggctgcaatggcatggtcacactc 1056

Qy   2617 ctgggaagcctgattggtctgagaatagccgattcgttgccttttccatgaaagatgaaa 2676
           || | |  || |||||||||| |  || |||||||||| ||  |  ||  || ||
Db   1057 ctagaactccagattggtctgaaacaagtcgattcgttgcattcacactggtcgacaaag 1116

Qy   2677 gacagggcgagatctatgtggccttcaacaccagccacttaccggccgttgttgagctcc 2736
             |||| ||  | ||| | ||||| ||| ||||||| || || |      ||   || |
Db   1117 tgaagggagaactatatattgcctttaacgccagccatttgcctgtaacgattacacttc 1176

Qy   2737 cagagcgcgcagggcgccggtgggaaccggtggtggacacaggcaagccagcaccatacg 2796
           ||||  |  | ||     | ||| | ||| | |||||||||||||| ||||||||| |
Db   1177 cagataggcctggttatagatggcagccgtttgtggacacaggcaaaccagcaccatttg 1236

Qy   2797 acttcctcaccgacgacttacctgatcgcgctctcaccatacacagttctcgcatttcc 2856
           ||||| | || |||||| | |||||  | |       |   | || |  || ||||| |
Db   1237 acttcttgacagacgacgttcctgagagagagacagcagccaaacaatattctcatttcc 1296

Qy   2857 tctactccaacctctaccccatgctcagctactcatcggtcatcctagtattgcgccctg 2916
           |  || | ||||  || || |||||||| || |||||  | || ||  || |     |||
Db   1297 tggacgcgaaccagtatccgatgctcagttattcatccattattcttttactatcatctg 1356

Qy   2917 atgtt 2921
             || |
Db   1357 ctgat 1361


RESULT   9
AAX27062
ID   AAX27062 standard; DNA; 2634 BP.
XX
AC   AAX27062;
XX
DT   21-MAY-1999  (first entry)
XX
DE   S. tuberosum isoamylase clone 9 coding sequence.
XX
KW   Isoamylase clone; potato; branched polysaccharide synthesis; amylopectin;
KW   starch modification; debranching activity; ss.
XX
OS   Solanum tuberosum.
XX
PN   WO9906575-A1.
XX
PD   11-FEB-1999.
XX
PF   30-JUL-1998;   98WO-GB02280.
XX
PR   31-JUL-1997;   97GB-0016185.
XX
PA   (PLAN-) PLANT BIOSCIENCE LTD.
XX
PI   Bustos Guillen R,  Edwards EA,  Martin CR,  Smith AM;
XX
DR   WPI; 1999-153803/13.
DR   P-PSDB; AAY00870.
XX
PT   New nucleic acid from potato encoding starch debranching enzyme with
PT   isoamylase activity - and related vectors, transformed cells and
PT   plants, proteins and antibodies, used to generate starch with
PT   altered properties, for use in foods, as thickeners etc.
XX
PS   Claim 3; Fig 3; 78pp; English.
XX
CC   This sequence encodes a Solanum tuberosum isoamylase of the invention.
CC   Fragments of the isoamylase coding sequence are used to identify and
CC   clone isoamylases from other plant species, by standard hybridisation
CC   or amplification methods. Expression of the isoamylase DNA in host cells
CC   is used to produce the isoamylase which is used: (i) to raise antibodies
CC   (Ab) (for identification, isolation and localisation of isoamylases); and
CC   (ii) for synthesis of branched polysaccharides. The DNA, proteins
CC   and Ab (or related peptides) are used to alter the quality and
CC   quantity of polysaccharides in a host cells, specifically to alter
CC   branching in amylopectin. The modified starches produced are useful in
CC   human or animal foods; as biodegradable plastic; as food or paint
CC   thickener; in starch-coated films, paper and textiles; in mining
CC   explosives; in pharmaceuticals and glues. Manipulation of debranching
CC   activity in a plant allows control of starch properties, e.g. increased
CC   gel strength; formation of paste rather than gel, changed physical
CC   characteristics etc.
XX
SQ   Sequence 2634 BP; 854 A; 496 C; 573 G; 711 T; 0 other;


  Query Match              5.0%;  Score 149.2;  DB 20;  Length 2634;
  Best Local Similarity   51.0%;  Pred. No. 2.1e-22;
  Matches  633;  Conservative    0;  Mismatches  508;  Indels   99;  Gaps    8;

Qy    412 gaggttccccttgacccccgatgaatcggaccgggaacgtgtggcatgtcttcatcgaa 471
          ||  |  |  | || || | || ||| || || ||  || | |||||  | | |||  |
Db    434 gaattagcattggatccacagaagaaccgcacaggagacatatggcacatatgcattaag 493

Qy    472 ggcgagctgca---caacatgctttacgggtacaggttcgacggcacctttgctcctcac 528
          |    ||  ||        | ||||| || ||  |  | || ||  |      |   ||
Db    494 gagttgccccaaggtggtgtcctttatggttatcgcattgatggacctcgaaattggcat 553

Qy    529 tgcgggcactaccttgatgtttccaatgtcgtggtggatccttatgctaaggcagtgata 588
            |||||     |||||| |  || | |  |||| |||||||| || ||   | |  |
Db    554 gaagggcatcgatttgatgatagcattattttggttgatccttacgcaaaactaattgaa 613

Qy    589 agccgaggggagtatggtgttccagcgcgtggtaacaattgctggcctcagatggctggc 648
           | ||| | |  | ||| | |  |      |   |    | | |    |   |    ||
Db    614 ggtcgacgagtttttggagatgaaagcaataaaatgtgtagatttttggaacttatgat 673
```

```
Qy    649 atgatccctcttccatatagcacgtttgattgggaaggcgacctacctctaagatatcct 708
Db    674 ttcaatagcttgccttttgactgg------ggagaaaattacaagcttccaaatataccc 727

Qy    709 caaaaggacctggtaatatatgagatgcacttgcgtggattcacgaagcatgattcaagc 768
Db    728 gagaaagatcttgttatatatgagatgaatgttcgtgcttttactgctgatgaaacaagt 787

Qy    769 a------atgtagaacatccgggtactttcattggagctgtgtcgaagcttgactatttg 822
Db    788 agtttggatcaagatcaacggggaagttaccttggcttaattgaaaagataccacatctt 847

Qy    823 aaggagcttggagttaattgtattgaattaatgccctgccatgagttcaacgagctggag 882
Db    848 ctcgagcttggtgtcaatgcagtagaattattgcctgtttttgagtttgatgaactggaa 907

Qy    883 tactcaacctcttcttc------------caagatgaacttttggggatattctaccata 930
Db    908 ttacaaaggcgacctaatccgagagatcacatgatcaatacatggggctactcaacaata 967

Qy    931 aacttcttttcaccaatgacaagatacacatcaggcgggataaaaaactgtgggcgtgat 990
Db    968 aacttttttgctccaatgagtcgatatgcaagttgtgg------tggcggacctgtccgt 1021

Qy    991 gccataaatgagttcaaaacttttgtaagagaggctcacaaacggggaattgaggtgatc 1050
Db   1022 gcttcctgggagttcaaagaaatggtcaaggccttgcatggtgctggaattgaggtcatc 1081

Qy   1051 ctggatgttgtcttcaaccatacagctgagggtaatgagaatggtcca---atattatca 1107
Db   1082 ttagatgttgtttataatcacacaaatgaagctgatgatgaaaacccatatacaacctca 1141

Qy   1108 tttaaggggtcgataatactacatactatatgcttg------cacccaagggagagttt 1161
Db   1142 ttccgaggaatagacaacaaggtttattacatggtagatttaaacaacaatgctcagctg 1201

Qy   1162 tataactattctggctgtgggaataccttcaactgtaatcatcctgtggttcgtcaattc 1221
Db   1202 ctgaatttcgctggatgtggaaatacttttaactgcaaccatcccacagtcatggaactt 1261

Qy   1222 attgtagattgtttaagatactgggtgacggaaatgcatgttgatggttttcgttttgat 1281
Db   1262 atacttgaaagcttaagacactgggtcaccgagtatcatgtcgatggatttcgctttgat 1321

Qy   1282 cttgcatccataatgaccagaggttccagtctgtgggatccagttaacgtgtatggagct 1341
Db   1322 cttgctagtgttctttgcagagg---------------------------------- 1344

Qy   1342 ccaatagaaggtgacatgatcacaacagggacacctcttgttactccaccacttattgac 1401
Db   1345 --------------------gacagatggtactcccattaatgctcccccccttgttaag 1384

Qy   1402 atgatcagcaatgacccaattcttggaggcgtcaagctcattgctgaagcatgggatgca 1461
Db   1385 gccatttccaaagatagtgtattgtcgaggtgcaaaattattgctgagccatgggattgt 1444

Qy   1462 ggaggcctctatcaagtaggtcaattccctcactggaatgtttggtctgagtggaatggg 1521
Db   1445 ggaggcctatatcttgttggaaagtttccgaactgggaccggtgggctgagtggaatggg 1504

Qy   1522 aagtaccgggacattgtgcgtcaattcattaaaggcactgatggatttgctggtggtttt 1581
Db   1505 aagtaccgcgatgacatcaggagatttataaagggcgatgctggcatgaaaggaaatttt 1564

Qy   1582 gccgaatgtctttgtggaagtccacacctataccaggtaa 1621
Db   1565 gcaacccgtatcgcaggttcagcggatctgtacagagtga 1604
```

RESULT 10
AAC81914/c
ID AAC81914 standard; DNA; 273254 BP.
XX
AC AAC81914;
XX
DT 27-FEB-2001 (first entry)
XX
DE Chlamydia pneumoniae genome DNA.
XX
KW Genome; diagnosis; vaccine; ds.
XX
OS Chlamydia pneumoniae.
XX
PN WO200027994-A2.
XX
PD 18-MAY-2000.
XX
PF 12-NOV-1999; 99WO-US26923.
XX
PR 12-NOV-1998; 98US-0108279.
PR 08-APR-1999; 99US-0128606.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Stephens R, Mitchell W, Kalman S, Davis R;
XX
DR WPI; 2000-376516/32.
XX
PT Isolated nucleic acid for use in diagnostic and analytical methods
PT encodes genomic sequence of Chlamydia pneumoniae -
XX
PS Claim 2; Page 128-320; 320pp; English.
XX
CC This invention describes a novel nucleic acid (N1) encoding a Chlamydia
CC pneumoniae protein (P1), given in the specification. The isolated nucleic
CC acid is useful for diagnostic and analytical methods, such as,
CC hybridization-based assays or amplification-based assays. The protein may
CC be used for diagnostic purposes, for their enzymatic or structural
CC activity, or as a vaccine. The invention also describes (1) a probe
CC comprising a hybridizing fragment of N1; (2) an isolated nucleic acid
CC (N2) that hybridizes under stringent conditions to N1; (3) an expression
CC cassette comprising N1 under the transcriptional regulation of a
CC transcriptional initiation region functional in an expression host, and a
CC transcriptional termination region; (4) a cell comprising an expression
CC cassette of (3) as part of an extrachromosomal element or integrated into
CC the genome of a host cell as a result of induction of the expression
CC cassette into the host cell, and the cellular progeny of the host cell;
CC (5) a method for producing a P1 comprising growing a cell of (4) where
CC the protein is expressed and isolating the protein free of other
CC proteins; (6) a purified polypeptide composition comprising at least 50
CC weight % of P1; and (7) a monoclonal antibody binding specifically to the
CC peptide of (6).
XX
SQ Sequence 273254 BP; 76423 A; 51054 C; 61965 G; 83812 T; 0 other;

Query Match 4.8%; Score 143; DB 21; Length 273254;
Best Local Similarity 55.3%; Pred. No. 2.2e-20;
Matches 359; Conservative 0; Mismatches 260; Indels 30; Gaps 3;

```
Qy    671 cgtttgattgggaaggcgacctacctctaagatatcctcaaaaggacctggtaatatatg 730
Db 170596 CATTTCCTTGGGATGGTGATCAGCCTCTGCATTTGCCGAAAGAAGAGATGATCATCTATG 170537

Qy    731 agatgcacttgcgtggattcacgaagcatgattcaagcaatgtagaacatccgggtactt 790
Db 170536 AGATGCATGTACGTTCCTTCACGCAATCTTCTTCATCTAGGGTTCATGCTCCGGGAACCT 170477

Qy    791 tcattggagctgtgtcgaagcttgactatttgaaggagcttggagttaattgtattgaat 850
Db 170476 TCCTAGGAATCATTGAAAAGATCGACCATCTGCATAAGCTGGGAATCAACGCTGTTGAAC 170417

Qy    851 taatgccctgccatgagttcaacgagctggag-----------tactcaacctcttctt 898
Db 170416 TCTTACCTATCTTTGAGTTCGATGAGACTGCGCATCCTTTTAGAAATTCGAAATTCCCTT 170357
```

```
Qy    899 ccaagatgaacttttggggatattctaccataaacttcttttcaccaatgacaagatataca 958
Db 170356 ATCTGTGCAATTATTGGGGTTATGCTCCCCTAAATTTCTTTTCTCCTTGCCGACGTTATG 170297

Qy    959 catcaggcgggataaaaaactgtgggcgtgatgccataaatgagttcaaaacttttgtaa 1018
Db 170296 CTTATGCC---------TCTGATCCTTGCGCTCCAAGTAGAGAGTTTAAAACTTTAGTAA 170246

Qy   1019 gagaggctcacaaacggggaattgaggtgatcctggatgttgtcttcaaccatacagctg 1078
Db 170245 AGACCTTGCATCAAGAAGGTATTGAGGTCATTCTTGATGTTGTTTTTAATCATACGGGCT 170186

Qy   1079 agggtaatgagaatggtccaatattatcatttaaggggtcgataatactacatactata 1138
Db 170185 TGC---------AAGGGACGACCTGCTCTTTGCCTTGGATAGACACTCCGAGCTATTATA 170135

Qy   1139 tgcttgcacccaagggagagttttataactattctggctgtgggaataccttcaactgta 1198
Db 170134 TTTTAGATGCACAAGGTCACTTTACAAATTATTCAGGCTGTGGAAACACTCTCAATACAA 170075

Qy   1199 atcatcctgtggttcgtcaattcattgtagattgtttaagatactgggtgacggaaatgc 1258
Db 170074 ACCGCGCCCCCACGACCCAATGGATTCTCGACATCTTACGTTATTGGGTAGAAGAAATGC 170015

Qy   1259 atgttgatggttttcgttttgatcttgcatccataatgaccagaggttc 1307
Db 170014 ATGTCGATGGGTTCCGATTTGATCTTGCTTCTGTCTTTTCTCGTGGTCC 169966
```

RESULT 11
AAX91990
ID AAX91990 standard; DNA; 1230025 BP.
XX
AC AAX91990;
XX
DT 13-SEP-1999 (first entry)
XX
DE Nucleotide sequence of the complete genome of Chlamydia pneumoniae.
XX
KW Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;
KW sinusitis; purulent otitis media; erythema nodosum; pharyngitis;
KW vaccine; neutralising epitope; ss.
XX
OS Chlamydia pneumoniae.
XX
PN WO9927105-A2.
XX
PD 03-JUN-1999.
XX
PF 20-NOV-1998; 98WO-IB01890.
XX
PR 04-NOV-1998; 98US-0107078.
PR 21-NOV-1997; 97FR-0014673.
XX
PA (GEST) GENSET.
XX
PI Griffais R;
XX
DR WPI; 1999-357842/30.
XX
PT Genome sequence of Chlamydia pneumoniae
XX
PS Claim 1; Page 291-611; 1912pp; English.
XX
CC The present sequence represents the complete genome of Chlamydia
CC pneumoniae, and encodes proteins AAY34584-Y35879. C. pneumoniae causes
CC respiratory disease such as pneumonia and bronchitis and is thought
CC to be a contributing factor in heart disease, sarcoidosis, sinusitis,
CC purulent otitis media, erythema nodosum or pharyngitis. The polypeptides
CC encoded by the open reading frames of the C. pneumoniae genome (see
CC AAY34584-Y35879) can be used in immunogenic compositions as vaccines.
CC Vectors containing C. pneumoniae nucleotides sequences can also be
CC used as immunogenic compositions, especially where the vector directs
CC the expression of a neutralising epitope of C. pneumoniae.
XX
SQ Sequence 1230025 BP; 367213 A; 249833 C; 249013 G; 363589 T; 377 other;

Query Match 4.8%; Score 143; DB 20; Length 1230025;
Best Local Similarity 55.3%; Pred. No. 3.6e-20;
Matches 359; Conservative 0; Mismatches 260; Indels 30; Gaps 3;

```
Qy    671 cgtttgattgggaaggcgacctacctctaagatatcctcaaaaggacctggtaatatatg 730
Db 446195 catttccttgggatggtgatcagcctctgcatttgccgaaagaagagatgatcatctatg 446254

Qy    731 agatgcacttgcgtggattcacgaagcatgattcaagcaatgtagaacatccgggtactt 790
Db 446255 agatgcatgtacgttccttcacgcaatcttcttcatctagggttcatgctccgggaacct 446314

Qy    791 tcattggagctgtgtcgaagcttgactatttgaaggagcttggagttaattgtattgaat 850
Db 446315 tcctaggaatcattgaaaagatcgaccatctgcataagctgggaatcaacgctgttgaac 446374

Qy    851 taatgccctgccatgagttcaacgagctggag----------tactcaacctcttctt 898
Db 446375 tcttacctatctttgagttcgatgagactgcgcatccttttagaaattcgaaattccctt 446434

Qy    899 ccaagatgaacttttggggatattctaccataaacttcttttcaccaatgacaagatataca 958
Db 446435 atctgtgcaattattggggttatgctcccctaaatttcttttctccttgccgacgttatg 446494

Qy    959 catcaggcgggataaaaaactgtgggcgtgatgccataaatgagttcaaaacttttgtaa 1018
Db 446495 cttatgcc---------tctgatccttgcgctccaagtagagagtttaaaacttttagtaa 446545

Qy   1019 gagaggctcacaaacggggaattgaggtgatcctggatgttgtcttcaaccatacagctg 1078
Db 446546 agaccttgcatcaagaaggtattgaggtcattcttgatgttgtttttaatcatacgggct 446605

Qy   1079 agggtaatgagaatggtccaatattatcatttaaggggtcgataatactacatactata 1138
Db 446606 tgc---------aagggacgacctgctctttgccttggatagacactccgagctattata 446656

Qy   1139 tgcttgcacccaagggagagttttataactattctggctgtgggaataccttcaactgta 1198
Db 446657 ttttagatgcacaaggtcactttacaaattattcaggctgtggaaacactctcaatacaa 446716

Qy   1199 atcatcctgtggttcgtcaattcattgtagattgtttaagatactgggtgacggaaatgc 1258
Db 446717 accgcgcccccacgacccaatggattctcgacatcttacgttattgggtagaagaaatgc 446776

Qy   1259 atgttgatggttttcgttttgatcttgcatccataatgaccagaggttc 1307
Db 446777 atgtcgatgggttccgatttgatcttgcttctgtcttttctcgtggtcc 446825
```

RESULT 12
AAX27067
ID AAX27067 standard; DNA; 1389 BP.
XX
AC AAX27067;
XX
DT 21-MAY-1999 (first entry)
XX
DE Original S. tuberosum isoamylase clone 9 coding sequence.
XX
KW Isoamylase clone; potato; branched polysaccharide synthesis; amylopectin;
KW starch modification; debranching activity; ss.
XX
OS Solanum tuberosum.
XX
PN WO9906575-A1.
XX
PD 11-FEB-1999.
XX

```
PF   30-JUL-1998;   98WO-GB02280.
XX
PR   31-JUL-1997;   97GB-0016185.
XX
PA   (PLAN-) PLANT BIOSCIENCE LTD.
XX
PI   Bustos Guillen R,  Edwards EA,  Martin CR,  Smith AM;
XX
DR   WPI; 1999-153803/13.
DR   P-PSDB; AAY00873.
XX
PT   New nucleic acid from potato encoding starch debranching enzyme with
PT   isoamylase activity - and related vectors, transformed cells and
PT   plants, proteins and antibodies, used to generate starch with
PT   altered properties, for use in foods, as thickeners etc.
XX
PS   Disclosure; Fig 12; 78pp; English.
XX
CC   This sequence represents the original isolated clone encoding a
CC   Solanum tuberosum isoamylase of the invention.
CC   Fragments of the isoamylase coding sequence are used to identify and
CC   clone isoamylases from other plant species, by standard hybridisation
CC   or amplification methods. Expression of the isoamylase DNA in host cells
CC   is used to produce the isoamylase which is used: (i) to raise antibodies
CC   (Ab) (for identification, isolation and localisation of isoamylases); and
CC   (ii) for synthesis of branched polysaccharides. The DNA, proteins
CC   and Ab (or related peptides) are used to alter the quality and
CC   quantity of polysaccharides in a host cells, specifically to alter
CC   branching in amylopectin. The modified starches produced are useful in
CC   human or animal foods; as biodegradable plastic; as food or paint
CC   thickener; in starch-coated films, paper and textiles; in mining
CC   explosives; in pharmaceuticals and glues. Manipulation of debranching
CC   activity in a plant allows control of starch properties, e.g. increased
CC   gel strength; formation of paste rather than gel, changed physical
CC   characteristics etc.
XX
SQ   Sequence 1389 BP; 437 A; 249 C; 320 G; 383 T; 0 other;


  Query Match             4.3%;  Score 129.6;  DB 20;  Length 1389;
  Best Local Similarity   53.0%;  Pred. No. 2.8e-18;
  Matches  357;  Conservative    0;  Mismatches  299;  Indels   18;  Gaps    3;

Qy   2155 ggaaaccttggcacagtatcaactttgtatgtgcacatgatggatttacactggctgatt 2214
           | || || |  |||||| ||||||| ||   ||| |||||||| |||||  ||  |||
Db    474 gaaagccgtaccacagtgtcaacttcgtgattgcccatgatggctttaccttgtatgacc 533

Qy   2215 tggtaacatataataagaagtacaatttaccaaatggggagaacaacagagatggagaaa 2274
           | ||  |||| ||||| ||| |||||    | || || ||      ||  ||||||   |
Db    534 ttgtttcatacaataataagcacaatgatgcgaacggtgaaggtggcaatgatggatgca 593

Qy   2275 atcacaatcttagctggaattgtggggaggaaggagaattcgcaagattgtctgtcaaaa 2334
           || ||||  | || ||||||||||    ||||| |||    ||        | | ||
Db    594 atgacaacttcagttggaattgtggaattgaaggtgaaacttcagatgcaaatattaacg 653

Qy   2335 gattgaggaagaggcagatgcgcaatttctttgtttgtctcatggtttctcaaggagttc 2394
           | || |    |||| |||  |||||  | |   || |||||||||| |||  |
Db    654 cactgcgttcacggcaaatgaaaaattttcatttggcactgatggtttctcagggaacac 713

Qy   2395 caatgttctacatgggtgatgaatatggccacacaaaaggggcaacaacaatacatact 2454
           ||||| |   ||||| ||||| ||||| || ||       || || || || | ||
Db    714 caatgatgcttatgggggatgagtatgggcatacccgctatggaaataataacagttatg 773

Qy   2455 gccatgattcttatgtcaattattttcgctgggataaaaaagaacaat---actctgagt 2511
           | |||||| |    |||| |||| | |||   ||  ||| |   |  |||
Db    774 gacatgataccgccatcaacaatttccagtggggacaattggaagcaaggaagaatgatc 833

Qy   2512 tgcaccgattctgctgcctcatgaccaaattccgcaaggagtgcgagggtcttggccttg 2571
             | | ||||  | |  ||||  || || ||   |    | | | |||| |   |
Db    834 acttcaggttcttttccaagatgataaagtttcgactgtcccacaatgttcttagaaagg 893

Qy   2572 aggactttccaacggccaaacggctgcagtggcatggtcatcagcctgggaagcctgatt 2631
           |  ||||      |   ||     |   |||| |   | | |     ||
Db    894 aaaacttcattgagaagaacgacattacctggctcgaggacaactggtacaa-------- 945

Qy   2632 ggtctgagaatagccgattcgttgccttttccatgaaagatgaaagacagggcgagatct 2691
             |||  | ||  ||||| |||| |||    |  | ||||  |    || || || |
Db    946 ----tgaagagagtagattccttgcatttatgctccatgatgggaatggaggagatattt 1001

Qy   2692 atgtggccttcaacaccagccacttaccggccgttgttgagctcccagagcgcgcagggc 2751
           |  |||| || ||  |   ||||||  |   |      |   | ||    |   || |
Db   1002 acttggcatttaatgcacaccacttttccatcaaaacagcaataccttcaccaccacgaa 1061

Qy   2752 ---gccggtgggaaccggtggtggacacaggcaagccagcaccatacgacttcctcaccg 2808
             |  | ||| | |  ||||||||||      |    ||||| | || ||  | |  |
Db   1062 atagaagttggtaccgagtggtggacactaatctgaagtcaccagatgattttgttattg 1121

Qy   2809 acgacttacctgat 2822
           | |   |  ||| |
Db   1122 agggagtgtctggt 1135


RESULT  13
AAZ01425
ID   AAZ01425 standard; DNA; 1038602 BP.
XX
AC   AAZ01425;
XX
DT   07-OCT-1999  (first entry)
XX
DE   Complete genome sequence of Chlamydia trachomatis.
XX
KW   Vaccine; eye disease; conventional trachoma; nonendemic trachoma;
KW   paratrachoma; inclusion conjunctivitis; genital disease; perihepatitis;
KW   nongonococcal uretritis; epidymitis; cervicitis; salpingitis;
KW   bartholinitis; pneumopathy; venereal lymphogranulomatosis; ss.
XX
OS   Chlamydia trachomatis.
XX
PN   WO9928475-A2.
XX
PD   10-JUN-1999.
XX
PF   27-NOV-1998;   98WO-IB01939.
XX
PR   04-NOV-1998;   98US-0107077.
PR   28-NOV-1997;   97FR-0015041.
PR   17-DEC-1997;   97FR-0016034.
XX
PA   (GEST ) GENSET.
XX
PI   Griffais R;
XX
DR   WPI; 1999-371125/31.
XX
PT   Genome sequence of Chlamydia trachomatis
XX
PS   Claim 1; Page 373-656; 1755pp; English.
XX
CC   The present sequence represents the complete genome of Chlamydia
CC   trachomatis. Open reading frames (ORFs) of the genome encode
CC   polypeptides AAY36754-Y37949. The polypeptides can be used as vaccines
CC   against Chlamydia trachomatis. Antisense and ribozyme sequences can also
CC   be used to control growth of the microorganism. Chlamydia trachomatis is
CC   responsible for a large number of diseases, e.g. eye diseases such as
CC   conventional trachoma, nonendemic trachoma, paratrachoma, and inclusion
CC   conjunctivitis; genital diseases such as nongonococcal uretritis,
CC   epidymitis, cervicitis, salpingitis, perihepatitis, bartholinitis;
CC   pneumopathy in breast feeding infants; and venereal
CC   lymphogranulomatosis. The polypeptides of the invention may be of use in
CC   treating these diseases.
XX
SQ   Sequence 1038602 BP; 304265 A; 214645 C; 214259 G; 305001 T; 432 other;


  Query Match             4.0%;  Score 120.8;  DB 20;  Length 1038602;
```

```
  Best Local Similarity   56.6%;  Pred. No. 2.1e-15;
  Matches  224;  Conservative    0;  Mismatches  172;  Indels    0;  Gaps    0;

Qy   2155 ggaaaccttggcacagtatcaactttgtatgtgcacatgatggatttacactggctgatt 2214
          ||||||||  | ||   || |||| | | ||  | |||||||| || |||||    |||
Db 424349 ggaaacctacgaactcgattaactatatctgctctcatgatggcttcacactctacgata 424408

Qy   2215 tggtaacatataataagaagtacaatttaccaaatggggagaacaacagagatggagaaa 2274
            ||  | |||||  | ||| |||||  |   ||||| ||  ||||  | |||||
Db 424409 ctgttgcctataacgataagcacaatgaagagaatggtgaatacaatcgtgatgggactt 424468

Qy   2275 atcacaatcttagctggaattgtggggaggaaggagaattcgcaagattgtctgtcaaaa 2334
              ||   |||||  || | |||    |||||||||    ||       |  |
Db 424469 cagcaaactatagctataactttggctgcgaaggagaaacgacagatcccaccatttgcg 424528

Qy   2335 gattgaggaagaggcagatgcgcaatttctttgtttgtctcatggtttctcaaggagttc 2394
            |  |  |  | || |||   || |||||| ||  |||| |  | |||||||||| |||
Db 424529 ctctacgtgaacgccaaatgaaaaacttctttcttgctctttttatctcaaggaattc 424588

Qy   2395 caatgttctacatgggtgatgaatatggccacacaaaaggggcaacaacaatacatact 2454
          | ||| |  |    || |||||||||| ||||||      || || || |||   |  |
Db 424589 ccatgatacaatccggagatgaatatgggcacacagcttatggaaataataatcactggt 424648

Qy   2455 gccatgattcttatgtcaattatttttcgctgggataaaaaagaacaatactctgagttgc 2514
          ||   ||  || |  ||||||| ||||  ||||||  |  ||   ||      ||  ||
Db 424649 gcttagacactaagatcaattacttttctttgggatcgattagctgaaaggaaagaactgt 424708

Qy   2515 accgattctgctgcctcatgaccaaattccgcaagg 2550
            ||||  ||||   | |      | |||||
Db 424709 tttctttcttatgccaagtcattgctctgcgcaaag 424744


RESULT  14
AAX76376
ID   AAX76376 standard; DNA; 2523 BP.
XX
AC   AAX76376;
XX
DT   04-AUG-1999  (first entry)
XX
DE   Zea mays partial sugary 1 genomic sequence.
XX
KW   Zea mays; SU1; starch debranching enzyme; maize; endosperm; corn;
KW   bacterial isoamylase; alpha-amylase; starch hydrolytic enzyme;
KW   sugary 1; ds.
XX
OS   Zea mays.
XX
PN   US5912413-A.
XX
PD   15-JUN-1999.
XX
PF   24-MAR-1995;   95US-0410784.
XX
PR   24-MAR-1995;   95US-0410784.
XX
PA   (IOWA ) UNIV IOWA STATE RES FOUND INC.
XX
PI   James MG,  Myers AM;
XX
DR   WPI; 1999-357234/30.
XX
PT   SU1 starch debranching enzyme
XX
PS   Claim 4; Fig 2; 39pp; English.
XX
CC   The present invention describes a starch debranching enzyme active in
CC   maize endosperm (Zea mays), designated SU1. SU1 protein is useful as
CC   a replacement for the bacterial and fungal enzymes currently used in
CC   the starch processing industry. The present sequence represents a
CC   Zea mays partial sugary 1 genomic sequence.
XX
SQ   Sequence 2523 BP; 591 A; 623 C; 614 G; 695 T; 0 other;

  Query Match             4.0%;  Score 120;  DB 20;  Length 2523;
  Best Local Similarity   80.1%;  Pred. No. 4e-16;
  Matches  141;  Conservative    0;  Mismatches   35;  Indels    0;  Gaps    0;

Qy     79 cggttgtcgaggcggcgacgaaggtagaggacgaggggggaggaggacgagccggtggcgg 138
          | || | |||||||| |  |  ||  |||||||| |  || || ||||||  |||||| |
Db    913 ccgtggccgaggcggtgcaggcggaggaggacgacgacgacgacgaggaggtggccg 972

Qy    139 aggacaggtacgcgctcggcggcgcgtgcagggtgctcgccggaatgcccgcgccgctgg 198
          |||| |||| |||||| |||||||||||| |||||||||| ||||||||||||||||| |
Db    973 aggagaggttcgcgctgggcggcgcgtgccgggtgctcgcgggaatgcccgcgccgctcg 1032

Qy    199 gcgccaccgcgctcgccggcggggtcaatttcgccgtctattccggcggagccacc 254
          |||||||||||||  |||||| |||| |||||||||| ||| |||    | ||
Db   1033 gcgccaccgcgctccgcggcggtgtcaacttcgccgtctactccagcgcccacccc 1088


RESULT  15
AAS59613
ID   AAS59613 standard; DNA; 22934 BP.
XX
AC   AAS59613;
XX
DT   13-FEB-2002  (first entry)
XX
DE   Propionibacterium acnes immunogenic protein encoding DNA #108.
XX
KW   SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW   uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW   inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW   dermatological; osteopathic; neuroprotectant; ds.
XX
OS   Propionibacterium acnes.
XX
PN   WO200181581-A2.
XX
PD   01-NOV-2001.
XX
PF   20-APR-2001; 2001WO-US12865.
XX
PR   21-APR-2000; 2000US-199047P.
PR   02-JUN-2000; 2000US-208841P.
PR   07-JUL-2000; 2000US-216747P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Skeiky YAW,  Persing DH,  Mitcham JL,  Wang SS,  Bhatia A;
PI   L'maisonneuve J,  Zhang Y,  Jen S,  Carter D;
XX
DR   WPI; 2001-616774/71.
XX
PT   Propionibacterium acnes polypeptides and nucleic acids useful for
PT   vaccinating against and diagnosing infections, especially useful for
PT   treating acne vulgaris -
XX
PS   Claim 1; SEQ ID No 108; 1069pp; English.
XX
CC   Sequences AAS59506-AAS59804 represent DNA molecules encoding
CC   Propionibacterium acnes immunogenic polypeptides. The proteins and their
CC   associated DNA sequences are used in the treatment, prevention and
CC   diagnosis of medical conditions caused by P. acnes. The disorders include
CC   SAPHO syndrome (synovitis, acne, pustulosis, hypertosis and
CC   osteomyelitis), uveitis and endophthalmitis. P. acnes is also involved
CC   in infections of bone, joints and the central nervous system, however it
CC   is particularly involved in the inflammatory lesions associated with acne
CC   vulgaris. A method for detecting the presence or absence of P. acnes in a
CC   patient comprises contacting a sample with a binding agent that binds to
CC   the proteins of the invention and determining the amount of bound protein
CC   in the sample. The polypeptides may be used as antigens in the production
CC   of antibodies specific for P. acnes proteins. These antibodies can be
CC   used to downregulate expression and activity of P. acnes polypeptides and
```

CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA). This sequence encodes the
CC polypeptides shown in AAU60884-AAU61133 and AAU67652-AAU67654.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 22934 BP; 4591 A; 7273 C; 6479 G; 4591 T; 0 other;

Query Match 3.9%; Score 116.2; DB 23; Length 22934;
Best Local Similarity 55.6%; Pred. No. 5.6e-15;
Matches 223; Conservative 0; Mismatches 178; Indels 0; Gaps 0;

```
Qy   2168 cagtatcaactttgtatgtgcacatgatggatttacactggctgatttggtaacatataa 2227
          |||  |||||||  |    || || || || || ||  ||  |||  |||| || || ||
Db  11355 cagcgtcaacttcattaccgctcacgacggcttcacgatgcgtgacctggtcacctacaa 11414

Qy   2228 taagaagtacaatttaccaaatggggagaacaacagagatggagaaaatcacaatcttag 2287
           | |||| | ||   |   || || ||| | |||   || ||    ||| |||| |
Db  11415 catgaagcataaccaaggcaacggcgaggataaccacgacgggtctaatgacaaccgggc 11474

Qy   2288 ctggaattgtggggaggaaggagaattcgcaagattgtctgtcaaaagattgaggaagag 2347
          |||||| |||||   ||| || ||     |      |||  ||   |  ||| || |  |
Db  11475 ctggaactgtggttgggagggcgagactaccgatcagtccatcgtcatgttgcggcatcg 11534

Qy   2348 gcagatgcgcaatttctttgtttgtctcatggtttctcaaggagttccaatgttctacat 2407
           ||| |||||||  |  | | |  | |  |  |  ||   || || ||||||  ||  |
Db  11535 tcaggtgcgcaacctggtcgctactttgctattagctaccggtgtaccaatgatcaccgc 11594

Qy   2408 gggtgatgaatatggccacacaaaaggggcaacaacaatacatactgccatgattctta 2467
           ||||| |||   || | |||  |   |||||| |||||| |||||||||| |||||·
Db  11595 cggtgacgaaatgggacgcacccagcagggcaataacaatgcatactgccaggattcccc 11654

Qy   2468 tgtcaattattttcgctgggataaaaaagaacaatactctgagttgcaccgattctgctg 2527
          | ||   |   |    |||||  |    |||  ||  || ||  ||  |    ||  |
Db  11655 tatctcgtggatagattgggaggatgccgaagcatggtccgatgtgaccgacctcgtcaa 11714

Qy   2528 cctcatgaccaaattccgcaaggagtgcgagggtcttggcc 2568
            | || ||   | | ||    |||  |   | |||| | |
Db  11715 gataattacggcactgcgtcgcgagcaccccgctcttcgtc 11755
```

Search completed: July 31, 2002, 12:11:05
Job time: 10939 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2002, 08:19:06 ; Search time 69.73 Seconds
(without alignments)
10557.344 Million cell updates/sec

Title: US-09-674-817A-1
Perfect score: 2997
Sequence: 1 ggtcggggccggcgccgcgc.....aaaaaaaaaaaaaaaaaaaa 2997

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1000	33.4	2712	2	US-08-410-784A-1	Sequence 1, Appli
	2	575.4	19.2	2133	4	US-09-187-124-1	Sequence 1, Appli
	3	120	4.0	2523	2	US-08-410-784A-3	Sequence 3, Appli
c	4	102.8	3.4	4411529	4	US-09-103-840A-1	Sequence 1, Appli
	5	80.6	2.7	2181	3	US-09-129-075-3	Sequence 3, Appli
	6	80.6	2.7	2181	4	US-09-346-237-12	Sequence 12, Appl
c	7	59	2.0	71989	4	US-09-443-501A-2	Sequence 2, Appli
	8	53.8	1.8	1028	4	US-08-118-200-1	Sequence 1, Appli
	9	53.8	1.8	1028	4	US-08-458-745-1	Sequence 1, Appli
	10	53.2	1.8	2625	6	5457037-4	Patent No. 5457037
	11	53.2	1.8	3336	6	5457037-1	Patent No. 5457037
	12	52.8	1.8	8051	2	US-08-576-626A-2	Sequence 2, Appli
	13	52.4	1.7	1315	4	US-09-721-822A-10	Sequence 10, Appl
	14	51.6	1.7	2253	6	5457037-2	Patent No. 5457037
	15	50.8	1.7	535	4	US-09-056-556-171	Sequence 171, App
	16	50.8	1.7	1069	4	US-09-372-422A-7	Sequence 7, Appli
c	17	50.8	1.7	1931	2	US-09-130-114-2	Sequence 2, Appli
	18	49.8	1.7	2550	6	5258287-23	Patent No. 5258287
	19	49.8	1.7	4895	3	US-09-053-866-1	Sequence 1, Appli
	20	49.6	1.7	3238	4	US-08-123-934A-5	Sequence 5, Appli
	21	49.6	1.7	3238	5	PCT-US94-10080-5	Sequence 5, Appli
	22	49.4	1.6	635	1	US-08-455-633A-35	Sequence 35, Appl
	23	49.4	1.6	635	1	US-08-416-336-5	Sequence 5, Appli
	24	49.4	1.6	635	2	US-08-456-460C-35	Sequence 35, Appl
	25	49.4	1.6	635	5	PCT-US94-05354-35	Sequence 35, Appl
	26	49.4	1.6	2674	4	US-09-817-180-1	Sequence 1, Appli
	27	49	1.6	1325	2	US-08-464-517-1	Sequence 1, Appli
	28	49	1.6	1325	2	US-08-246-361A-1	Sequence 1, Appli
	29	49	1.6	1325	3	US-08-463-772-1	Sequence 1, Appli
	30	49	1.6	1325	5	PCT-US93-05000-1	Sequence 1, Appli
c	31	48.8	1.6	152331	3	US-09-128-155-16	Sequence 16, Appl
	32	48.6	1.6	1029	3	US-08-911-853-1	Sequence 1, Appli
	33	48.6	1.6	1029	4	US-09-479-409-1	Sequence 1, Appli
	34	48.6	1.6	1029	4	US-09-479-453-1	Sequence 1, Appli
	35	48.6	1.6	1454	2	US-08-713-000-7	Sequence 7, Appli
	36	48.6	1.6	1454	2	US-08-975-316-7	Sequence 7, Appli
	37	48.6	1.6	1454	4	US-09-211-710-7	Sequence 7, Appli
	38	48.6	1.6	1474	2	US-08-975-316-71	Sequence 71, Appl
	39	48.6	1.6	4377	3	US-08-911-853-28	Sequence 28, Appl
	40	48.6	1.6	4377	4	US-09-479-409-28	Sequence 28, Appl
	41	48.6	1.6	4377	4	US-09-479-453-28	Sequence 28, Appl
	42	48.4	1.6	1096	4	US-09-461-697-26	Sequence 26, Appl
	43	48.2	1.6	1813	5	PCT-US94-12883-3	Sequence 3, Appli
	44	48.2	1.6	2218	2	US-08-985-090-4	Sequence 4, Appli
	45	48.2	1.6	2218	3	US-09-165-543-31	Sequence 31, Appl

ALIGNMENTS

```
RESULT 1
US-08-410-784A-1
; Sequence 1, Application US/08410784A
; Patent No. 5912413
;  GENERAL INFORMATION:
;    APPLICANT:  MYERS, ALAN M.
;    APPLICANT:  JAMES, MARTHA G.
;    TITLE OF INVENTION:  ISOLATION OF SU1, A STARCH DEBRANCHING
;    TITLE OF INVENTION:  ENZYME, THE PRODUCT OF THE MAIZE GENE
;    TITLE OF INVENTION:  SUGARY 1
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Weingarten, Schurgin, Gagnebin and Hayes LLP
;      STREET:  Ten Post Office Square
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/410,784A
;      FILING DATE:  24-MAR-1995
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Heine, Ph.D., Holliday C
;      REGISTRATION NUMBER:  34,346
;      REFERENCE/DOCKET NUMBER:  ISU-002XX
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-542-2290
;      TELEFAX:  617-451-0313
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2712 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:
;    ORIGINAL SOURCE:
;    FEATURE:
```

```
;      NAME/KEY:  Coding Sequence
;      LOCATION:  1...2454
;      OTHER INFORMATION:
US-08-410-784A-1


  Query Match             33.4%;  Score 1000;  DB 2;  Length 2712;
  Best Local Similarity   65.7%;  Pred. No. 2.5e-225;
  Matches 1883;  Conservative    0;  Mismatches  350;  Indels  633;  Gaps    4;

Qy      79 cggttgtcgaggcggcgacgaaggtagaggacgaggggaggaggacgagccggtggcgg 138
           | || | ||||||||  |   |  ||  ||||||||  |  || || ||||||  |||||| |
Db     233 CCGTGGCCGAGGCGGTGCAGGCGGAGGAGGACGACGACGACGACGAGGAGGTGGCCG 292

Qy     139 aggacaggtacgcgctcggcggcgcgtgcagggtgctcgccggaatgcccgcgccgctgg 198
           ||||  |||| ||||||  |||||||||||| ||||||||||  |||||||||||||||| |
Db     293 AGGAGAGGTTCGCGCTGGGCGGCGCGTGCCGGGTGCTCGCGGGAATGCCCGCGCCGCTCG 352

Qy     199 gcgccaccgcgctcgccggcggggtcaatttcgccgtctattccggcggagccaccgccg 258
           ||||||||||||||  ||||||  |||||  |||||||||||  |||  ||||  |||  ||||||
Db     353 GCGCCACCGCGCTCCGCGGCGGTGTCAACTTCGCCGTCTACTCCAGCGGTGCCTCCGCCG 412

Qy     259 cggcgctctgcctcttcacgccagaagatctcaaggcggtggggttgcctcccgagtaga 318
           ||  ||||  ||||||||  |  ||  |  ||  ||||||||
Db     413 CGTCGCTGAGCCTCTTCGCTCCCGGCGACCTCAAGGC---------------------- 449

Qy     319 gttcatcagctttgcgtgcgccgcgcgccccttttttgggcctgcaatttaagttttgta 378

Db     450 ------------------------------------------------------------ 449

Qy     379 ctggggcaaatgctgcaggatagggtgaccgaggaggttccccttgaccccctgatgaat 438
                           |||||||||||||||||||| ||||| || ||||||  |  ||
Db     450 -----------------GGATAGGGTGACCGAGGAGGTGCCCCTCGATCCCCTGCTCAAC 492

Qy     439 cggaccgggaacgtgtggcatgtcttcatcgaagg---cgagctgcacaacatgctttac 495
           ||  ||  ||  ||||||||||| ||  ||||||  |  ||    ||||||||||  ||||||  |  |
Db     493 CGAACGGGAAACGTGTGGCACGTGTTCATCCACGGGGACGAGCTGCACGGCATGCTCTGC 552

Qy     496 gggtacaggttcgacggcacctttgctcctcactgcgggcactaccttgatgtttccaat 555
           ||  ||||||||||| |||    ||  ||  |||  |  ||||  ||  |||    |||||  |||||
Db     553 GGATACAGGTTCGATGGCGTGTTCGCCCCTGAGCGCGGACAGTACTACGATGTGTCCAAC 612

Qy     556 gtcgtggtggatccttatgctaaggcagtgataagccgaggggagtatggtgttccagcg 615
           ||  ||||||||||| ||  ||||||||||| ||||||||||  ||  ||||||||  ||  |||
Db     613 GTTGTGGTGGATCCATACGCTAAGGCAGTGGTAAGCCGAGGTGAATATGGTGTGCCTGCG 672

Qy     616 cgtggtaacaattgctggcctcagatggctggcatgatccctcttccatatagcacgttt 675
           |  ||||   |  |||  ||||||||  ||||||||  ||||||||||||||  ||||  |  ||||
Db     673 CCTGGTGGTAGTTGTTGGCCTCAAATGGCTGGTATGATCCCTCTTCCCTATAATAAGTTT 732

Qy     676 gattgggaaggcgacctacctctaagatatcctcaaaaggacctggtaatatatgagatg 735
           ||||||  ||||  ||||||||  ||  |  ||  |  |||  ||||||||  ||  ||||||||  |||
Db     733 GATTGGCAAGGTGACCTACCCCTTGGGTACCATCAGAAGGACCTTGTCATATATGAAATG 792

Qy     736 cacttgcgtggattcacgaagcatgattcaagcaatgtagaacatccgggtactttcatt 795
           ||  ||||||||||||||  |||||  |  ||||||||  |  ||||  ||  ||  ||||  ||||
Db     793 CATTTGCGTGGATTCACAAAGCACAACTCAAGCAAGACAAAACACCCAGGAACTTACATT 852

Qy     796 ggagctgtgtcgaagcttgactatttgaaggagcttggagttaattgtattgaattaatg 855
           ||  ||||||||  ||||||||||  ||  |  |||||  ||||||||  ||  |||||  ||  |||||
Db     853 GGTGCTGTGTCAAAGCTTGACCATCTAAAGGAACTTGGAGTGAACTGTATAGAGCTAATG 912

Qy     856 ccctgccatgagttcaacgagctggagtactcaacctcttcttccaagatgaacttttgg 915
           ||||||||||||||||  |||||  |||||||  |  |||  |||||  ||||||||||  |||
Db     913 CCCTGCCATGAGTTCAATGAGCTAGAGTACTTCAGCTCCTCTTCGAAGATGAACTTCTGG 972

Qy     916 ggatattctaccataaacttcttttcaccaatgacaagatacacatcaggcgggataaaa 975
           ||||||||  ||  |||||  ||  ||  |||||||||  |||||||  |  |||  |  ||  ||||  .|
Db     973 GGATATTCCACAATAAATTTTTTCTCACCAATGGCAAGATATTCTTCAAGTGGCATAAGA 1032

Qy     976 aactgtgggcgtgatgccataaatgagttcaaaacttttgtaagagaggctcacaaacgg 1035
             |||  |||  |||  ||||||||||||  ||  |||  ||||||||||  |||||  |||||||||
Db    1033 GACTCTGGATGTGGTGCCATAAATGAATTTAAAGCTTTTGTAAGGGAGGCCCACAAACGG 1092

Qy    1036 ggaattgaggtgatcctggatgttgtcttcaaccatacagctgagggtaatgagaatggt 1095
           ||||||||||||||  ||||||||||||||||  ||||||||||  ||||||||||  ||
Db    1093 GGAATTGAGGTGATCATGGATGTTGTCTTCAATCATACAGCTGAAGGTAATGAGAAAGGC 1152

Qy    1096 ccaatattatcatttaaggggtcgataatactacatactatatgcttgcacccaaggga 1155
           ||||||||||  ||||  ||||  |  |||||||  |||||||||  ||||||||||  ||||||
Db    1153 CCAATATTATCCTTTAGGGGGATAGATAATAGTACATACTACATGCTTGCACCTAAGGGA 1212

Qy    1156 gagttttataactattctggctgtgggaataccttcaactgtaatcatcctgtggttcgt 1215
           ||||||||||  ||||||||  |||||  ||||||||||  ||||||||||||||  ||  |||
Db    1213 GAGTTTTATAATTATTCTGGTTGTGGAAATACCTTCAATTGTAATCATCCTGTAGTCCGT 1272

Qy    1216 caattcattgtagattgtttaagatactgggtgacggaaatgcatgttgatggttttcgt 1275
           ||||  ||  ||  |||||  ||  ||||||||||  ||  ||||||||||||||||||||||||
Db    1273 GAATTTATAGTGGATTGCTTGAGATACTGGGTAACAGAAATGCATGTTGATGGTTTTCGT 1332

Qy    1276 tttgatcttgcatccataatgaccagaggttccagtctgtgggatccagttaacgtgtat 1335
           |||||  ||||||||  |||  ||||||||||  |  ||||||  ||||||||||||||  ||||||
Db    1333 TTTGACCTTGCATCTATACTGACCAGAGGATGCAGTCTATGGGATCCAGTTAATGTGTAT 1392

Qy    1336 ggagctccaatagaaggtgacatgatcacaacagggacacctcttgttactccaccactt 1395
           |||  ||||||  ||||||||||||||  ||  ||||||||||||||||||  |  |||||||||
Db    1393 GGAAGTCCAATGGAAGGTGACATGATTACGACAGGGACACCTCTTGTTGCCCCACCACTT 1452

Qy    1396 attgacatgatcagcaatgacccaattcttggaggcgtcaagctcattgctgaagcatgg 1455
           ||||||||||  ||||||||||||||||||||  ||||||||||||||||||||||||
Db    1453 ATTGACATGATTAGCAATGACCCAATTCTTGGAAATGTCAAGCTCATTGCTGAAGCATGG 1512

Qy    1456 gatgcaggaggcctctatcaagtaggtcaattccctcactggaatgtttggtctgagtgg 1515
           ||||||||||  ||||||||||  ||||||  ||  ||||||||||  ||||||||  ||||||
Db    1513 GATGCAGGAGGTCTCTATCAAGAAGGTCAGTTTCCTCACTGGAACGTTTGGTCAGAGTGG 1572

Qy    1516 aatgggaagtaccgggacattgtgcgtcaattcattaaaggcactgatggatttgctggt 1575
           |||||  |||||  ||  ||  |  ||||||||  |||||  ||||||||  ||||||||||||||||
Db    1573 AATGGAAAGTATCGCGATACCGTGCGTCAGTTCATCAAAGGCACAGATGGATTTGCTGGT 1632

Qy    1576 ggttttgccgaatgtctttgtggaagtccacacctataccaggtaagttgtggcaatact 1635
           |  ||||||  |||||  ||  ||||||||||||||  |||||
Db    1633 GCTTTTGCTGAATGCCTATGTGGAAGTCCACAGTTATAC---------------------- 1671

Qy    1636 tgtaaatgagttgagtgaatgtcacctggatttttttatatataccacatgatgatacaca 1695

Db    1672 ------------------------------------------------------------ 1671

Qy    1696 tctaaatatataacaatcatagtgtatgcatatgcatttggctaagaagtattagtgtat 1755

Db    1672 ------------------------------------------------------------ 1671

Qy    1756 acactagtgctatatataggttttaacacccaacttgccaatgaaggaacatagggcttt 1815

Db    1672 ------------------------------------------------------------ 1671

Qy    1816 ctagttatcttatttatttgtccggtgaataatccactgaaaaattccagccatgtcatt 1875

Db    1672 ------------------------------------------------------------ 1671

Qy    1876 ttttaggggggagaagaaactatattgatttgccccccctaaaagaagccatctcagaat 1935

Db    1672 ------------------------------------------------------------ 1671

Qy    1936 tcataggtaagttgctttttctgtaaagaaaggaaaacgacttcatactttctatcggtgc 1995

Db    1672 ------------------------------------------------------------ 1671

Qy    1996 taacttagctcgatgtatatttgtaagatgaatgccaaatttaatttgtcggataatttg 2055

Db    1672 ------------------------------------------------------------ 1671

Qy    2056 atctgttattcacaaatttctatttggtttctctagaaatcaaaccagtaacttgttatt 2115

Db    1672 ------------------------------------------------------------ 1671
```

```
Qy   2116 ggcactgcaacttcttattgattaatcaggcaggaggaaggaaaccttggcacagtatca 2175
                                 ||||||||  ||  |||||  ||||||||||||||||
Db   1672 -------------------------CAGGCAGGGGGGAGGAAGCCTTGGCACAGTATCG 1705

Qy   2176 actttgtatgtgcacatgatggatttacactggctgatttggtaacatataataagaagt 2235
           ||||||||||||||| ||||||||||||||||||||||||| |||||  ||||  ||||
Db   1706 GCTTTGTATGTGCACACGATGGATTTACACTGGCTGATTTGGTCACATACAATAGCAAGT 1765

Qy   2236 acaatttaccaaatggggagaacaacagagatggagaaaatcacaatcttagctggaatt 2295
          ||||  ||  |||||||  |||  ||  |||||||||  ||||||||  ||||||||||||||||
Db   1766 ACAACTTGTCAAATGGTGAGGACTTCAGAGATGGGGAAAATCATAATCTTAGCTGGAATT 1825

Qy   2296 gtggggaggaaggagaattcgcaagattgtctgtcaaaagattgaggaagaggcagatgc 2355
          ||||||||||||||||||  |||||  ||||  |||  ||||||  ||||||||||  ||||
Db   1826 GTGGGGAGGAAGGAGAATTTGCAAGTCTGTCAGTCCGAAGATTAAGGAAGAGGCAAATGC 1885

Qy   2356 gcaatttctttgtttgtctcatggtttctcaaggagttccaatgttctacatgggtgatg 2415
          ||||||||||||||||||  ||||||||||  ||||||||||||||||||||||||  ||||
Db   1886 GCAATTTCTTTGTTTGTCTTATGGTTTCTCAGGGAGTTCCAATGTTCTACATGGGCGATG 1945

Qy   2416 aatatggccacacaaaaggggcaacaacaatacatactgccatgattcttatgtcaatt 2475
          |||||||  ||||||||  ||  ||  ||||||||||  ||||||||||    ||||||||||
Db   1946 AATATGGTCACACAAAGGGAGGGAACAACAATACGTACTGCCATGACCATTATGTCAATT 2005

Qy   2476 attttcgctgggataaaa---aagaacaatactctgagttgcaccgattctgctgcctca 2532
          ||||  ||  ||||||||  |   |||||||||  ||||||  |||  ||  ||||||||  |  ||||
Db   2006 ATTTCCGTTGGGATAAGAAGGAAGAACAATCCTCTGATTTGTACAGATTCTGCCGTCTCA 2065

Qy   2533 tgaccaaattccgcaaggagtgcgagggtcttggccttgaggactttccaacggccaaac 2592
          |||||  ||||||||||  ||  ||  ||   ||||||||||||||||||  ||  ||  |  |||
Db   2066 TGACCGAATTCCGCAAAGAATGTGAATCTCTTGGCCTTGAGGACTTCCCGACTTCAGAAC 2125

Qy   2593 ggctgcagtggcatggtcatcagcctgggaagcctgattggtctgagaatagccgattcg 2652
          ||  ||  |  |||||  ||||||||||  ||||||||||  |||||  |||    ||||||||||
Db   2126 GGTTGAAATGGCACGGTCATCAGCCCGGGAAGCCTGACTGGTCAGAGGCAAGCCGATTCG 2185

Qy   2653 ttgccttttccatgaaagatgaaagacagggcgagatctatgtggccttcaacaccagcc 2712
          |||||||  |||||||  ||  ||||   |  ||||||||||  ||||||||||||||||  |
Db   2186 TTGCCTTCACCATGAAGGACGAAACCAAAGGCGAGATCTACGTGGCCTTCAACACCAGTC 2245

Qy   2713 acttaccggccgttgttgagctcccagagcgcgcagggcgccggtgggaaccggtggtgg 2772
          ||  |  ||||  |||||||  |||  ||||||||||  |  |||  |||  |||||  ||||||||||
Db   2246 ACCTTCCGGTGGTTGTTGGGCTTCCAGAGCGCTCTGGGTTCCGATGGGAGCCGGTGGTGG 2305

Qy   2773 acacaggcaagccagcaccatacgacttcctcaccgacgacttacctgatcgcgctctca 2832
          ||||  ||||||   |||||||  ||||||||||||||||  |  |  |  ||  |||||  |||  |||
Db   2306 ACACCGGCAAGGAGGCACCATATGACTTCCTCACCGATGGCCTGCCAGATCGTGCTGTCA 2365

Qy   2833 ccatacaccagttctcgcatttcctctactccaacctctaccccatgctcagctactcat 2892
          ||  |  ||||||||||  |||||||||  |||||||  |||||  ||  ||||||||||||||  |
Db   2366 CCGTCTACCAGTTCTCTCATTTCCTCAACTCCAATCTCTATCCTATGCTCAGCTACTCCT 2425

Qy   2893 cggtcatcctagtattgcgccctgatgtttgagagaccaatatata 2938
          |  |||||||  ||||||||||||||||  |||  |||   |  |||  |
Db   2426 CCATCATCCTTGTATTGCGCCCTGATGTCTGAAAGAAGCAGATACA 2471


RESULT   2
US-09-187-124-1
; Sequence 1, Application US/09187124A
; Patent No. 6255563
; GENERAL INFORMATION:
;  APPLICANT: Emmermann, Michael
;  APPLICANT: Kossmann, Jens
;  TITLE OF INVENTION: NUCLEIC ACID MOLECULES CODING FOR DEBRANCHING ENZYMES
;  TITLE OF INVENTION: FROM POTATO
;  FILE REFERENCE: GFB8
;  CURRENT APPLICATION NUMBER: US/09/187,124A
;  CURRENT FILING DATE: 1998-11-05
;  EARLIER APPLICATION NUMBER: PCT/EP97/02292
;  EARLIER FILING DATE: 1997-05-06
;  EARLIER APPLICATION NUMBER: DE 196 18 125.9
;  EARLIER FILING DATE: 1996-05-06
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 2133
;   TYPE: DNA
;   ORGANISM: Solanum tuberosum
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (2)..(1819)
;   OTHER INFORMATION: Clone: Iso5
US-09-187-124-1


  Query Match             19.2%;  Score 575.4;  DB 4;  Length 2133;
  Best Local Similarity   69.8%;  Pred. No. 9.5e-126;
  Matches  799;  Conservative    0;  Mismatches  331;  Indels   15;  Gaps    1;

Qy     600 gtatggtgttccagcgcgtggtaacaattgctggcctcagatggctggcatgatcccctct 659
           | ||   |  | || ||  |   |  |||| |||||  |  ||||| ||||||  | |||
Db       1 gaattcggcacgagggccagaggatgattgttggccccccaatggcaggcatggtaccttc 60

Qy     660 tccatatagcacgtttgattgggaaggcgacctacctctaagatatcctcaaaaggacct 719
           | | | |     ||||||||||||||| || |||   || |  | ||| || |  || ||
Db      61 tgcttctgatcagtttgattgggaaggagatctattactgaagtttccacagagagatct 120

Qy     720 ggtaatatatgagatgcacttgcgtggattcacgaagcatgattcaagcaatgtagaaca 779
           |||||  |||||  |||||  | |||||||| || || |||||  || ||  |   | || |
Db     121 tgtaatctatgaaatgcatgttcgtggatttacaaatcatgagtcgagtgaaacaaaata 180

Qy     780 tccgggtactttcattggagctgtgtcgaagcttgactatttgaaggagcttggagttaa 839
           |||  |||||||  |  ||||  |  ||||  |||  |||||  |  |||||||  |||||  || ||
Db     181 tcctggtacttaccttggtgttgtggagaaacttgatcacttgaaggaacttggtgtcaa 240

Qy     840 ttgtattgaattaatgccctgccatgagttcaacgagctggagtactcaacctcttcttc 899
           |||||  ||  ||||||||||  ||  ||||||||  ||||||||||||||  |  |  |  ||
Db     241 ctgtatagagctaatgccctgtcacgagttcaatgagctggagtactatagttataactc 300

Qy     900 ---------------caagatgaacttttggggatattctaccataaacttcttttcacc 944
                          ||||  |  ||||||||||  ||||||||  |  ||  ||||||||  ||
Db     301 tgtattgggcgactacaagtttaacttttggggctattctactgtcaatttcttttctcc 360

Qy     945 aatgacaagatacacatcaggcgggataaaaaactgtgggcgtgatgccataaatgagtt 1004
           ||||  |||||||  |  ||  |  ||  |||  ||  ||  ||  |  |  |||  |||||  ||  ||
Db     361 aatgggaagatactcgtctgctggtctaagtaattgcggcctcggtgcaataaacgaatt 420

Qy    1005 caaaacttttgtaagagaggctcacaaacggggaattgaggtgatcctggatgttgtctt 1064
           ||   |  ||||  |  ||  ||  ||  |||||  ||||  |||  ||||||||||  ||
Db     421 taagtatcttgtcaaggaagcacataaacgtggaatcgaggttatcatggatgttgtttt 480

Qy    1065 caaccatacagctgagggtaatgagaatggtccaatattatcatttaaggggtcgataa 1124
           |||  ||  ||  |||||  ||  |||||  ||||||||  |||  |||||||||  ||  |  ||  ||
Db     481 caatcacactgctgaaggaaatgaaaatggtcccatactatcatttagaggcattgacaa 540

Qy    1125 tactacatactatatgcttgcacccaagggagagttttataactattctggctgtgggaa 1184
           |  |   |  ||||  |||  ||  ||  |||||  ||  |||||  |||||  ||  ||  |||||  ||
Db     541 cagtgtgttttatacgctagctcctaagggtgaattttacaactactcaggatgtggaaa 600

Qy    1185 taccttcaactgtaatcatcctgtggttcgtcaattcattgtagattgtttaagatactg 1244
           ||||||||||||||||  ||||  |  ||  ||||||||  ||  ||  |||||  ||  ||||  ||
Db     601 taccttcaactgtaataatcccattgtacgtcaatttatagtggattgcttgagatattg 660

Qy    1245 ggtgacggaaatgcatgttgatggttttcgttttgatcttgcatccataatgaccagagg 1304
           |||  ||  ||||||||  ||  |||||  ||  ||  ||||||||||  ||  ||  |  ||  |||  |
Db     661 ggttaccgaaatgcacgtagatggcttccgctttgatcttgcttctatccttacaagaag 720

Qy    1305 ttccagtctgtgggatccagttaacgtgtatggagctccaatagaaggtgacatgatcac 1364
           |  |||   ||||  ||  |  ||  ||  ||  ||||||  |  ||||  ||  ||||||||||||||
Db     721 tagcagctcgtggaatgctgtaaatgtctatggaaattcaattgacggtgacatgatcac 780

Qy    1365 aacagggacacctcttgttactccaccacttattgacatgatcagcaatgacccaattct 1424
           |||||  ||  |||||    |  ||||||  |  ||||  ||||  ||||||||  |||||  ||
Db     781 cacaggcactcctctcacaagcccaccattgattgatatgattagcaatgatccaatact 840
```

```
Qy   1425 tggaggcgtcaagctcattgctgaagcatgggatgcaggaggcctctatcaagtaggtca 1484
          | | || || ||||| || ||||||||||||||||   |||||||| || ||||| ||
Db    841 tagtggagtaaagcttatagctgaagcatgggattgtggaggcctttaccaagttggcat 900

Qy   1485 attccctcactggaatgtttggtctgagtggaatgggaagtaccgggacattgtgcgtca 1544
            || || ||||||  | |·|||||| |||||||| || |||||||| ||||| || |||||
Db    901 gtttccgcactggggtatctggtcggagtggaacggaaagtaccgtgacatggtacgtca 960

Qy   1545 attcattaaaggcactgatggatttgctggtggttttgccgaatgtctttgtggaagtcc 1604
           ||||| |||||||||||||| ||| |||| | |||||| ||||| |||||||||||| ||
Db    961 gttcatcaaaggcactgatgggttttctggggcttttgctgaatgcctttgtggaagccc 1020

Qy   1605 acacctataccaggtaagttgtggcaatacttgtaaatgagttgagtgaatgtcacctgg 1664
          | | |||||||||  | |  |  | ||  | || ||  |  |  | |   |||   |
Db   1021 aaatctataccagaaaggaggaagaaaaccatggaacagtataaatttcgtgtgtgccca 1080

Qy   1665 atttttatatataccacatgatgatacacatctaaatatataacaatcatagtgtatgc 1724
             | | |   |    |    | |   |||| ||  ||| |   |   | |   |||
Db   1081 cgatggttttactttggctgatttagtgacatacaacaataaacacaatttggcaaatgg 1140

Qy   1725 atatg 1729
          | | |
Db   1141 agagg 1145


RESULT   3
US-08-410-784A-3
; Sequence 3, Application US/08410784A
; Patent No. 5912413
;  GENERAL INFORMATION:
;    APPLICANT:  MYERS, ALAN M.
;    APPLICANT:  JAMES, MARTHA G.
;    TITLE OF INVENTION:  ISOLATION OF SU1, A STARCH DEBRANCHING
;    TITLE OF INVENTION:  ENZYME, THE PRODUCT OF THE MAIZE GENE
;    TITLE OF INVENTION:  SUGARY 1
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Weingarten, Schurgin, Gagnebin and Hayes LLP
;      STREET:  Ten Post Office Square
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/410,784A
;      FILING DATE:  24-MAR-1995
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Heine, Ph.D., Holliday C
;      REGISTRATION NUMBER:  34,346
;      REFERENCE/DOCKET NUMBER:  ISU-002XX
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-542-2290
;      TELEFAX:  617-451-0313
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2523 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Genomic DNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:
;    ORIGINAL SOURCE:
US-08-410-784A-3


  Query Match             4.0%;  Score 120;  DB 2;  Length 2523;
  Best Local Similarity   80.1%;  Pred. No. 7.4e-19;
  Matches  141;  Conservative    0;  Mismatches   35;  Indels    0;  Gaps    0;

Qy     79 cggttgtcgaggcggcgacgaaggtagaggacgaggggggaggaggacgagccggtggcgg 138
          | || | |||||||| |  |  ||  |||||||| |  || || ||||||  |||||| |
Db    913 CCGTGGCCGAGGCGGTGCAGGCGGAGGAGGACGACGACGACGACGAGGAGGTGGCCG 972

Qy    139 aggacaggtacgcgctcggcggcgcgtgcagggtgctcgccggaatgcccgcgccgctgg 198
          |||| |||| |||||| |||||||||||| |||||||||| |||||||||||||||| |
Db    973 AGGAGAGGTTCGCGCTGGGCGGCGCGTGCCGGGTGCTCGCGGGAATGCCCGCGCCGCTCG 1032

Qy    199 gcgccaccgcgctcgccggcggggtcaatttcgccgtctattccggcggagccacc 254
          |||||||||||||| |||||| ||||| |||||||||| ||| |||    | ||
Db   1033 GCGCCACCGCGCTCCGCGGCGGTGTCAACTTCGCCGTCTACTCCAGCGCCCACCCC 1088


RESULT   4
US-09-103-840A-1/c
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
;  APPLICANT: FLEISCHMAN, Robert D.
;  APPLICANT:  WHITE, Owen R.
;  APPLICANT:  FRASER, Claire M.
;  APPLICANT:  VENTER, John C.
;  TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
;  TITLE OF INVENTION:  TUBERCULOSIS
;  FILE REFERENCE: 24366-20007.00
;  CURRENT APPLICATION NUMBER: US/09/103,840A
;  CURRENT FILING DATE:  1998-06-24
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 4411529
;   TYPE: DNA
;   ORGANISM: Mycobacterium tuberculosis
;   OTHER INFORMATION: H37Rv
US-09-103-840A-1


  Query Match             3.4%;  Score 102.8;  DB 4;  Length 4411529;
  Best Local Similarity   52.6%;  Pred. No. 2e-13;
  Matches  224;  Conservative    0;  Mismatches  202;  Indels    0;  Gaps    0;

Qy   2143 aggcaggaggaaggaaaccttggcacagtatcaactttgtatgtgcacatgatggattta 2202
          | |||   ||  |    ||  |   ||||||||| || ||    || || || || || |
Db 1770242 AAGCAACCGGCCGCCGGCCCAGTGCCAGTATCAATTTCGTCACCGCCCACGACGGGTTCA 1770183

Qy   2203 cactggctgatttggtaacatataataagaagtacaatttaccaaatggggagaacaaca 2262
          ||||    ||  ||||  | || ||  | ||| ||||     | ||||| |||||||||
Db 1770182 CACTCAACGACCTGGTCTCGTACAACGACAAGCACAACGAGGCCAATGGCGAGAACAACC 1770123

Qy   2263 gagatggagaaaatcacaatcttagctggaattgtggggaggaaggagaattcgcaagat 2322
          | || || ||||   |||| |     ||||| || || |  || ||      ||
Db 1770122 GCGACGGGGAAAGCTACAACCGATCGTGGAACTGCGGTGTCGAGGGCCCCACCGATGACC 1770063

Qy   2323 tgtctgtcaaaagattgaggaagaggcagatgcgcaatttctttgtttgtctcatggttt 2382
                ||       || |     | |||||||||||  | |  |     || |||||
Db 1770062 CCGACATCTTGGCGCTGCGTGCCCGCCAGATGCGCAACATGTGGGCCACGCTTATGGTCA 1770003

Qy   2383 ctcaaggagttccaatgttctacatgggtgatgaatatggccacacaaaaggggcaaca 2442
            || ||    || ||| ||  |   || || ||   ||| | |||  ||   |||||||
Db 1770002 GCCAGGGCACGCCGATGATCGCCCACGGCGACGAGATTGGGCGCACCCAATACGGCAACA 1769943

Qy   2443 acaatacatactgccatgattcttatgtcaattattttcgctgggataaaaaagaacaat 2502
          ||||    |||||||| || ||  |  |   ||   |   |||          ||  |
Db 1769942 ACAACGTCTACTGCCAGGACTCCGAATTATCTTGGATGGATTGGTCATTGGTGGACAAGA 1769883
```

```
Qy   2503 actctgagttgcaccgattctgctgcctcatgaccaaattccgcaaggagtgcgagggtc 2562
          |  | ||  |||     |||    ||     ||| |  || ||||||  |   | |||
Db 1769882 ATGCCGATCTGCTAGCTTTCGCACGCAAGGCGACGACCTTGCGCAAGAACCACAAGGTGT 1769823

Qy   2563 ttggcc 2568
          || |||
Db 1769822 TTCGCC 1769817


RESULT   5
US-09-129-075-3
; Sequence 3, Application US/09129075
; Patent No. 6087149
; GENERAL INFORMATION:
;  APPLICANT: Tsutsumi, No. 6087149iko
;  APPLICANT: Bisgard-Frantzen, Henrik
;  APPLICANT: Svendsen, Allan
;  TITLE OF INVENTION: Starch Conversion Process
;  FILE REFERENCE: 5262.204-US
;  CURRENT APPLICATION NUMBER: US/09/129,075
;  CURRENT FILING DATE: 1998-08-04
;  EARLIER APPLICATION NUMBER: 0787/97
;  EARLIER FILING DATE: 1997-07-02
;  EARLIER APPLICATION NUMBER: 60/055,867
;  EARLIER FILING DATE: 1997-08-13
;  EARLIER APPLICATION NUMBER: PCT/DK98/00304
;  EARLIER FILING DATE: 1998-07-02
;  NUMBER OF SEQ ID NOS: 16
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3
;   LENGTH: 2181
;   TYPE: DNA
;   ORGANISM: Rhodothermus marinus
US-09-129-075-3


  Query Match             2.7%;  Score 80.6;  DB 3;  Length 2181;
  Best Local Similarity   49.6%;  Pred. No. 1.2e-09;
  Matches  206;  Conservative    0;  Mismatches  209;  Indels    0;  Gaps    0;

Qy   2172 atcaactttgtatgtgcacatgatggatttacactggctgatttggtaacatataataag 2231
          |||||||| ||    || || || || || || |||| || |||| |  || |  ||
Db   1378 atcaacttcgtcacggcgcacgacggcttcacgctggaagacctggtcagctacacgaaa 1437

Qy   2232 aagtacaatttaccaaatggggagaacaacagagatggagaaaatcacaatcttagctgg 2291
          ||| ||||   | | ||| |||   |||| | || ||     |  | ||   ||| |
Db   1438 aagcacaacgaagcgaatctggaaggcaaccgggacggcatggacgaaaactacagcacg 1497

Qy   2292 aattgtggggaggaaggagaattcgcaagattgtctgtcaaaagattgaggaagaggcag 2351
          || || |||| ||| |||              ||| ||       |   || |   ||
Db   1498 aactgcggggtggagggacccacgcaggatccgtccgtgctggcctgccgggaagcgctc 1557

Qy   2352 atgcgcaatttctttgtttgtctcatggtttctcaaggagttccaatgttctacatgggt 2411
          | |||||   |  |       ||| |  | || || || || || ||| |       ||
Db   1558 aagcgcagcctgatcagcacgctctttctctcgcagggcgtgcccatgctgctgggcggc 1617

Qy   2412 gatgaatatggccacacaaaaggggcaacaacaatacatactgccatgattcttatgtc 2471
          || ||       | |||  |    ||||||||||  | || ||||| ||     |  ||
Db   1618 gacgagctgtcgcgcacgcagcacggcaacaacaacgcctattgccaggacaacgagatc 1677

Qy   2472 aattattttcgctgggataaaaaagaacaatactctgagttgcaccgattctgctgcctc 2531
          |  |  |    |||| |    |   |    |||| |     |||    |||
Db   1678 agctggtacaactggcagctcgacacgcgcaagcagcagtttctggagttcgtgcgccag 1737

Qy   2532 atgaccaaattccgcaaggagtgcgagggtcttggccttgaggactttccaacgg 2586
          | || |   || |||||| ||     | |  |  | |      | ||||  || |
Db   1738 acgatctggtttcgcaagcagcatcggagcttccggcgccgccatttttctgaccg 1792


RESULT   6
US-09-346-237-12
; Sequence 12, Application US/09346237A
; Patent No. 6265197
; GENERAL INFORMATION:
;  APPLICANT: Bisgard-Frantzen, Henrik
;  APPLICANT: Svendsen, Allan
;  TITLE OF INVENTION: Starch Debranching Enzymes
;  FILE REFERENCE: 5629.200-US
;  CURRENT APPLICATION NUMBER: US/09/346,237A
;  CURRENT FILING DATE: 1999-07-01
;  EARLIER APPLICATION NUMBER: PA 1998 00868
;  EARLIER FILING DATE: 1998-07-02
;  EARLIER APPLICATION NUMBER: 60/094,353
;  EARLIER FILING DATE: 1998-07-28
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 12
;   LENGTH: 2181
;   TYPE: DNA
;   ORGANISM: Rhodotermus marinus DSM 4252
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(2181)
;   OTHER INFORMATION: Isoamylase
US-09-346-237-12


  Query Match             2.7%;  Score 80.6;  DB 4;  Length 2181;
  Best Local Similarity   49.6%;  Pred. No. 1.2e-09;
  Matches  206;  Conservative    0;  Mismatches  209;  Indels    0;  Gaps    0;

Qy   2172 atcaactttgtatgtgcacatgatggatttacactggctgatttggtaacatataataag 2231
          |||||||| ||    || || || || || || |||| || |||| |  || |  ||
Db   1378 atcaacttcgtcacggcgcacgacggcttcacgctggaagacctggtcagctacacgaaa 1437

Qy   2232 aagtacaatttaccaaatggggagaacaacagagatggagaaaatcacaatcttagctgg 2291
          ||| ||||   | | ||| |||   |||| | || ||     |  | ||   ||| |
Db   1438 aagcacaacgaagcgaatctggaaggcaaccgggacggcatggacgaaaactacagcacg 1497

Qy   2292 aattgtggggaggaaggagaattcgcaagattgtctgtcaaaagattgaggaagaggcag 2351
          || || |||| ||| |||              ||| ||       |   || |   ||
Db   1498 aactgcggggtggagggacccacgcaggatccgtccgtgctggcctgccgggaagcgctc 1557

Qy   2352 atgcgcaatttctttgtttgtctcatggtttctcaaggagttccaatgttctacatgggt 2411
          | |||||   |  |       ||| |  | || || || || || ||| |       ||
Db   1558 aagcgcagcctgatcagcacgctctttctctcgcagggcgtgcccatgctgctgggcggc 1617

Qy   2412 gatgaatatggccacacaaaaggggcaacaacaatacatactgccatgattcttatgtc 2471
          || ||       | |||  |    ||||||||||  | || ||||| ||     |  ||
Db   1618 gacgagctgtcgcgcacgcagcacggcaacaacaacgcctattgccaggacaacgagatc 1677

Qy   2472 aattattttcgctgggataaaaaagaacaatactctgagttgcaccgattctgctgcctc 2531
          |  |  |    |||| |    |   |    |||| |     |||    |||
Db   1678 agctggtacaactggcagctcgacacgcgcaagcagcagtttctggagttcgtgcgccag 1737

Qy   2532 atgaccaaattccgcaaggagtgcgagggtcttggccttgaggactttccaacgg 2586
          | || |   || |||||| ||     | |  |  | |      | ||||  || |
Db   1738 acgatctggtttcgcaagcagcatcggagcttccggcgccgccatttttctgaccg 1792


RESULT   7
US-09-443-501A-2/c
; Sequence 2, Application US/09443501A
; Patent No. 6303342
; GENERAL INFORMATION:
;  APPLICANT: Kosan Biosciences, Inc.
;  APPLICANT:  Julien, Bryan
;  APPLICANT:  Katz, Leonard
;  APPLICANT:  Khosla, Chaitan
;  APPLICANT:  Tang, Li
;  APPLICANT:  Ziermann, Rainer
;  TITLE OF INVENTION: Recombinant Methods and Materials for Producing
;  TITLE OF INVENTION:  Epothilone and Epothilone Derivatives
;  FILE REFERENCE: 30062-20031.00
;  CURRENT APPLICATION NUMBER: US/09/443,501A
;  CURRENT FILING DATE:  1999-11-19
```

```
;  PRIOR APPLICATION NUMBER: US 60/130,560
;  PRIOR FILING DATE: 1999-04-22
;  PRIOR APPLICATION NUMBER: US 60/122,620
;  PRIOR FILING DATE: 1999-03-03
;  PRIOR APPLICATION NUMBER: US 60/119,386
;  PRIOR FILING DATE: 1999-02-10
;  PRIOR APPLICATION NUMBER: US 60/109,401
;  PRIOR FILING DATE: 1998-11-20
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 71989
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic construct
US-09-443-501A-2


  Query Match             2.0%;  Score 59;  DB 4;  Length 71989;
  Best Local Similarity   49.6%;  Pred. No. 0.00064;
  Matches  181;  Conservative    0;  Mismatches  180;  Indels    4;  Gaps    1;

Qy       2 gtcggggccggcgccgcgcctgcgacggtggcgacccaatgcgacg----gcggggaagg 57
           |||||| ||  |  | | |  ||  ||||| ||   |||||| ||     |  | |   |
Db     985 GTCGGGTCCTTCTTCAGGGCGTCGTAGGTGGTGAGGGAATGCGGCGTGACGACGAGGGCG 926

Qy      58 gggtcggcgaggtgtgcgccgcggttgtcgaggcggcgacgaaggtagaggacgaggggg 117
           | ||   ||  | | | | || || |   ||   |  |  | |||| | |  ||  | ||
Db     925 GCGTGCTCGCCGCGGGTGACGGGGATTGAGACTGGAGGCGGCAGGTGGCGCTCGGAGCGG 866

Qy     118 aggaggacgagccggtggcggaggacaggtacgcgctcggcggcgcgtgcagggtgctcg 177
            || || ||| |   |||||||  | ||   | | |   || | ||||  || |   ||||
Db     865 CGGCGGTCGATCACCTGGCGGACGGCACCGATGTGGATGGTGTCGCGCTCAAGCACCTCG 806

Qy     178 ccggaatgcccgcgccgctgggcgccaccgcgctcgccggcggggtcaatttcgccgtct 237
           |       | |     ||| |||| | ||      || ||     | |  |      | | |
Db     805 ACCAGGGCCTCTTCGAGCTCGGCGGCGCCCACGGCGTCGAGCAGTTGCAGAAGCCGGGCG 746

Qy     238 attccggcggagccaccgccgcggcgctctgcctcttcacgccagaagatctcaaggcgg 297
           ||  ||  |  | |     ||| ||||| ||  |    |||  |||   ||    || |
Db     745 ATCGCGCTGCCGACGTTATCGCCGCGCTCGGCGACGATGCGCAGGAATGCCTGGCTGCTG 686

Qy     298 tggggttgcctcccgagtagagttcatcagctttgcgtgcgccgcgcgccccttttttgg 357
           |||     | | ||||  ||      | ||  |||  |||      || | | |  | |
Db     685 CGGGCGGCGCGCGCGAGGCGATCAAGGCCGCGGTGCTCGCGGGCGCGGCGCTTCTCGTCG 626

Qy     358 gcctg 362
           || |
Db     625 ACCAG 621


RESULT   8
US-08-118-200-1
; Sequence 1, Application US/08118200
; Patent No. 6197500
;  GENERAL INFORMATION:
;    APPLICANT:  SUTHERLAND, Grant R
;    APPLICANT:  RICHARDS, Robert I
;    APPLICANT:  SCHLESSINGER, David
;    APPLICANT:  NAGARAJA, Ramaiah
;    APPLICANT:  KREMER, Eric J
;    APPLICANT:  YU, Sui
;    APPLICANT:  BAKER, Elizabeth
;    APPLICANT:  MULLEY, John C
;    APPLICANT:  MANDEL, Jean-Louis
;    APPLICANT:  PRITCHARD, Melanie April
;    APPLICANT:  LYNCH, Michael
;    TITLE OF INVENTION:  DNA SEQUENCES RELATED TO ISOLATED
;    TITLE OF INVENTION:  FRAGILE X SYNDROME
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BURNS, DOANE, SWECKER & MATHIS
;      STREET:  P.O. Box 1404
;      CITY:  Alexandria
;      STATE:  Virginia
;      COUNTRY:  United States
;      ZIP:  22313-1404
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/118,200
;      FILING DATE:  09-SEP-1993
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/802,650
;      FILING DATE:  05-DEC-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/672,232
;      FILING DATE:  20-MAR-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/638,518
;      FILING DATE:  04-JAN-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/966,517
;      FILING DATE:  23-DEC-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Crane-Feury, Sharon E
;      REGISTRATION NUMBER:  36,113
;      REFERENCE/DOCKET NUMBER:  020160-164
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 836-6620
;      TELEFAX:  (703) 836-2021
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1028 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-118-200-1


  Query Match             1.8%;  Score 53.8;  DB 4;  Length 1028;
  Best Local Similarity   57.4%;  Pred. No. 0.0017;
  Matches   97;  Conservative    0;  Mismatches   72;  Indels    0;  Gaps    0;

Qy       3 tcggggccggcgccgcgcctgcgacggtggcgacccaatgcgacggcggggaaggggtc 62
           || | || |||| || |||  || ||   |||  |    ||| |||||| |  || ||
Db     327 TCTGAGCGGGCGGCGGGCCGACGGCGAGCGCGGGCGGCGGCGGCGGCGGCGGCGGCG 386

Qy      63 ggcgaggtgtgcgccgcggttgtcgaggcggcgacgaaggtagaggacgaggggaggag 122
           |  | || |   || ||||  || | |||||||  ||  ||  | ||  | || || || |
Db     387 GCGGCGGCGGCGGCGGCGGCGGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 446

Qy     123 gacgagccggtggcggaggacaggtacgcgctcggcggcgcgtgcaggg 171
           |  | | ||| |||||  ||    |   || |   ||  |   | | |||
Db     447 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCCCGGAGCCACCTCTTCGGGG 495


RESULT   9
US-08-458-745-1
; Sequence 1, Application US/08458745
; Patent No. 6242576
;  GENERAL INFORMATION:
;    APPLICANT:  SUTHERLAND, Grant R
;    APPLICANT:  RICHARDS, Robert I
;    APPLICANT:  SCHLESSINGER, David
;    APPLICANT:  NAGARAJA, Ramaiah
;    APPLICANT:  KREMER, Eric J
;    APPLICANT:  YU, Sui
;    APPLICANT:  BAKER, Elizabeth
;    APPLICANT:  MULLEY, John C
```

```
;    APPLICANT:  MANDEL, Jean-Louis
;    APPLICANT:  PRITCHARD, Melanie April
;    APPLICANT:  LYNCH, Michael
;    TITLE OF INVENTION:  DNA SEQUENCES RELATED TO ISOLATED
;    TITLE OF INVENTION:  FRAGILE X SYNDROME
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BURNS, DOANE, SWECKER & MATHIS
;      STREET:  P.O. Box 1404
;      CITY:  Alexandria
;      STATE:  Virginia
;      COUNTRY:  United States
;      ZIP:  22313-1404
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/458,745
;      FILING DATE:  02-JUN-1995
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/118,200
;      FILING DATE:  09-SEP-1993
;      APPLICATION NUMBER:  US 07/802,650
;      FILING DATE:  05-DEC-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/672,232
;      FILING DATE:  20-MAR-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/638,518
;      FILING DATE:  04-JAN-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/966,517
;      FILING DATE:  23-DEC-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Crane-Feury, Sharon E
;      REGISTRATION NUMBER:  36,113
;      REFERENCE/DOCKET NUMBER:  020160-164
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 836-6620
;      TELEFAX:  (703) 836-2021
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1028 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-458-745-1


  Query Match             1.8%;  Score 53.8;  DB 4;  Length 1028;
  Best Local Similarity   57.4%;  Pred. No. 0.0017;
  Matches   97;  Conservative    0;  Mismatches   72;  Indels    0;  Gaps    0;

Qy       3 tcggggccggcgccgcgcctgcgacggtggcgacccaatgcgacggcggggaaggggtc 62
           || | || |||| || |||  || ||   |||  |    ||| |||||| |  || ||
Db     327 TCTGAGCGGGCGGCGGGCCGACGGCGAGCGCGGGCGGCGGCGGCGGCGGCGGCGGCG 386

Qy      63 ggcgaggtgtgcgccgcggttgtcgaggcggcgacgaaggtagaggacgagggggaggag 122
           |  | || |   || ||||  || | ||||||| ||  ||  | ||  | || || || |
Db     387 GCGGCGGCGGCGGCGGCGGCGGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 446

Qy     123 gacgagccggtggcggaggacaggtacgcgctcggcggcgcgtgcaggg 171
           |  | | ||| ||||| ||    |   || |   ||  |   | | |||
Db     447 GCGGCGGCGGCGGCGGCGGCGGCGGCCCGGAGCCACCTCTTCGGGG 495


RESULT  10
5457037-4
;Patent No. 5457037
;    APPLICANT: TOGNONI, ANGELO;CARRERA, PAOLO;CAMERINI, BARBARA;
;GALLI, GIULIANO;LUCCHESE, GIUSEPPE;GRANDI, GUIDO;DI GENNARO, CARLO
;    TITLE OF INVENTION: CLONING OF THE GENE CODING THE ISOAMYLASE
; ENZYME AND ITS USE IN THE PRODUCTION OF SAID ENZYME
;    NUMBER OF SEQUENCES: 7
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/1,797
;      FILING DATE: 08-JAN-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: 749,621
;      FILING DATE: 19-AUG-1991
;      APPLICATION NUMBER: 224,114
;      FILING DATE: 25-JUL-1988
;SEQ ID NO:4:
;      LENGTH: 2625
5457037-4


  Query Match             1.8%;  Score 53.2;  DB 6;  Length 2625;
  Best Local Similarity   49.4%;  Pred. No. 0.0035;
  Matches  315;  Conservative    0;  Mismatches  278;  Indels   45;  Gaps    5;

Qy     707 ctcaaaaggacctggtaatatatgagatgcacttgcgtggattcacgaagcatga---tt 763
           | || |||||    || || || ||| ||||  |||| || |||||  |||| ||    |
Db     830 cgcagaaggatgatgtgatctacgaggtgcatgtgcgcggcttcaccgagcaggacacct 889

Qy     764 caagcaatgtagaacatccgggtactttcattggagctgtgtcgaagcttgactatttga 823
           | | |  ||   |  |||  || || |     || || | |   |||      ||  |
Db     890 ctatccctgcgcagtatcgcggcacctattacggtgcagggctcaaggccagttacctcg 949

Qy     824 aggagcttggagttaattgtattgaattaatgccctgccatgagttcaacgagctggagt 883
             || || || |      | |||||  ||||    || ||     |  |    | |
Db     950 ccagcctgggcgtgaccgcggtggaattcctgccggtgcaggaaacgcagaatgatgcga 1009

Qy     884 ac---------tcaacctcttcttccaagatgaacttttggggatattctaccataaact 934
           ||          | |  ||   | ||||   |||||  |||||  ||    |||   ||||
Db    1010 acgatgtggttcccaattcagatgccaaccagaactactggggctacatgaccgagaact 1069

Qy     935 tcttttcaccaatgacaagatacacatcaggcgggataaaaaactgtgggcgtgatgcca 994
            ||| || ||         | ||  | |           || ||    |   | | | |||
Db    1070 acttctcgccggatcgccgctatgcct---------acaacaaggcggctggcggtccca 1120

Qy     995 taaatgagttcaaaacttttgtaagagaggctcacaaacggggaattgaggtgatcctgg 1054
             |||||| |  |  | ||    | |  ||||||    || ||  ||||   | |||
Db    1121 cggcggagttccaggcgatggtgcaggcgtttcacaacgcaggcatcaaggtctacatgg 1180

Qy    1055 atgttgtcttcaaccatacagctgaggg---------------------taatgagaatg 1093
           |||| |||| |||||| || || || ||                     |      | |
Db    1181 atgtggtctacaaccacaccgccgaaggcggcacctggaccagcagtgatcccaccacgg 1240

Qy    1094 gtccaatattatcatttaaggggtcgataatactacatactatatgcttgcacc---ca 1150
             | || |  || |      ||  | || ||||| || |||||   |||  |  |   ||
Db    1241 ccaccatttattcgtggcgcggcttggacaataccacgtactacgagctgacctcgggca 1300

Qy    1151 agggagagttttataactattctggctgtgggaataccttcaactgtaatcatcctgtgg 1210
           |   | | || ||  || |  | |||  |||    | ||||||     |  |  |.||||
Db    1301 accaatacttctacgacaacacgggcattggcgcgaacttcaatacgtacaacacggtgg 1360

Qy    1211 ttcgtcaattcattgtagattgtttaagatactgggtgacggaaatgcatgttgatggtt 1270
             |   | | || || || |   |    || |||| ||     |||   || ||||| |
Db    1361 cgcagaaccttatcgtcgactcgctggcgtattgggcgaacacgatgggcgtggatggct 1420

Qy    1271 ttcgttttgatcttgcatccataatgaccagaggttcc 1308
           |||| || || ||||| ||| |  |   ||   ||| |
Db    1421 ttcgcttcgaccttgcttccgtgctcggcaacagttgc 1458


RESULT  11
5457037-1
;Patent No. 5457037
;    APPLICANT: TOGNONI, ANGELO;CARRERA, PAOLO;CAMERINI, BARBARA;
;GALLI, GIULIANO;LUCCHESE, GIUSEPPE;GRANDI, GUIDO;DI GENNARO, CARLO
;    TITLE OF INVENTION: CLONING OF THE GENE CODING THE ISOAMYLASE
```

```
; ENZYME AND ITS USE IN THE PRODUCTION OF SAID ENZYME
;    NUMBER OF SEQUENCES: 7
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/1,797
;      FILING DATE: 08-JAN-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: 749,621
;      FILING DATE: 19-AUG-1991
;      APPLICATION NUMBER: 224,114
;      FILING DATE: 25-JUL-1988
;SEQ ID NO:1:
;      LENGTH: 3336
5457037-1


  Query Match             1.8%;  Score 53.2;  DB 6;  Length 3336;
  Best Local Similarity   49.4%;  Pred. No. 0.0039;
  Matches  315;  Conservative    0;  Mismatches  278;  Indels   45;  Gaps    5;

Qy     707 ctcaaaaggacctggtaatatatgagatgcacttgcgtggattcacgaagcatga---tt 763
           | || |||||    || || || ||| ||||  |||| || |||||  |||| ||    |
Db    1542 cgcagaaggatgatgtgatctacgaggtgcatgtgcgcggcttcaccgagcaggacacct 1601

Qy     764 caagcaatgtagaacatccgggtactttcattggagctgtgtcgaagcttgactatttga 823
           | | |  ||    |  |||  || || |     || || | |   |||       ||  |
Db    1602 ctatccctgcgcagtatcgcggcacctattacggtgcagggctcaaggccagttacctcg 1661

Qy     824 aggagcttggagttaattgtattgaattaatgccctgccatgagttcaacgagctggagt 883
             || || || |      | |||||  ||||    || ||      |  |     | |
Db    1662 ccagcctgggcgtgaccgcggtggaattcctgccggtgcaggaaacgcagaatgatgcga 1721

Qy     884 ac---------tcaacctcttcttccaagatgaacttttggggatattctaccataaact 934
           ||          | |  ||   | ||||   |||||  ||||| ||    |||   ||||
Db    1722 acgatgtggttcccaattcagatgccaaccagaactactggggctacatgaccgagaact 1781

Qy     935 tcttttcaccaatgacaagatacacatcaggcgggataaaaaactgtgggcgtgatgcca 994
            ||| || ||        | ||  | |           || ||    |   | | | |||
Db    1782 acttctcgccggatcgccgctatgcct---------acaacaaggcggctggcggtccca 1832

Qy     995 taaatgagttcaaaacttttgtaagagaggctcacaaacggggaattgaggtgatcctgg 1054
           ||||||  |  |  | ||    | |  ||||||    || ||  ||||   | |||
Db    1833 cggcggagttccaggcgatggtgcaggcgtttcacaacgcaggcatcaaggtctacatgg 1892

Qy    1055 atgttgtcttcaaccatacagctgaggg--------------------taatgagaatg 1093
           |||| |||| |||||| || || || ||                    |      | |
Db    1893 atgtggtctacaaccacaccgccgaaggcggcacctggaccagcagtgatcccaccacgg 1952

Qy    1094 gtccaatattatcatttaaggggtcgataatactacatactatatgcttgcacc---ca 1150
             | || |  || |     ||  | || ||||| || |||||  |||  |  |   ||
Db    1953 ccaccatttattcgtggcgcggcttggacaataccacgtactacgagctgacctcgggca 2012

Qy    1151 agggagagttttataactattctggctgtgggaatacttcaactgtaatcatcctgtgg 1210
           |   | | || ||  || |  | |||  |||    | ||||||     |  |  | ||||
Db    2013 accaatacttctacgacaacacgggcattggcgcgaacttcaatacgtacaacacggtgg 2072

Qy    1211 ttcgtcaattcattgtagattgtttaagatactgggtgacggaaatgcatgttgatggtt 1270
             |   | | || || || |   |    || |||| ||     |||   || |||||
Db    2073 cgcagaaccttatcgtcgactcgctggcgtattgggcgaacacgatgggcgtggatggct 2132

Qy    1271 ttcgttttgatcttgcatccataatgaccagaggttcc 1308
           |||| || || ||||| ||| |  |   ||   ||| |
Db    2133 ttcgcttcgaccttgcttccgtgctcggcaacagttgc 2170


RESULT  12
US-08-576-626A-2
; Sequence 2, Application US/08576626A
; Patent No. 5998194
;  GENERAL INFORMATION:
;    APPLICANT:  Summers, R.G.
;    APPLICANT:  Katz, L.
;    APPLICANT:  Donadio, S.
;    APPLICANT:  Staver, M.J.
;    TITLE OF INVENTION:  POLYKETIDE-ASSOCIATED SUGAR
;    TITLE OF INVENTION:  BIOSYNTHESIS GENES
;    NUMBER OF SEQUENCES:  60
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Abbott Laboratories
;      STREET:  100 Abbott Park Road
;      CITY:  Abbott Park
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60064-3500
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/576,626A
;      FILING DATE:  21-DEC-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Dianne Casuto
;      REGISTRATION NUMBER:  P-40,943
;      REFERENCE/DOCKET NUMBER:  5857.US.01
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (847) 938-3137
;      TELEFAX:  (847) 938-2623
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  8051 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-576-626A-2


  Query Match             1.8%;  Score 52.8;  DB 2;  Length 8051;
  Best Local Similarity   53.6%;  Pred. No. 0.0071;
  Matches  133;  Conservative    0;  Mismatches  112;  Indels    3;  Gaps    1;

Qy      74 cgccgcggttgtcgaggcggcgacgaaggtagaggacgaggggggaggaggacgagccggt 133
           || ||  ||   |   |    |  | | || ||||||| ||     | |  ||   | |
Db     694 CGGCGGCGTGCGCCGCGACCTGCTGCACGTCGAGGACGTGGCCACCGCGTTCGCCGCCGC 753

Qy     134 ggcggaggacaggtacgcgct---cggcggcgcgtgcagggtgctcgccggaatgcccgc 190
           |  |||| ||    |||||||   |||||||  |||| | ||  |||||      |||
Db     754 GCTGGAGCACCACGACGCGCTGGCCGGCGGCACGTGGGCGCTGGGCGCCGACCGATCCGA 813

Qy     191 gccgctgggcgccaccgcgctcgccggcggggtcaatttcgccgtctattccggcggagc 250
           |||||| |||| || |   |  |||| |   | ||   |||||    |  ||||| |  |
Db     814 GCCGCTCGGCGACATCTTCCGGGCCGTCTCCGGCAGCGTCGCCCGGCAGACCGGCAGCCC 873

Qy     251 caccgccgcggcgctctgcctcttcacgccagaagatctcaaggcggtggggttgcctcc 310
           | ||| ||  | | ||  | |   | |||| ||  |   | ||||     |  || |
Db     874 CGCCGTCGACGTGGTCACCGTGCCCGCGCCCGAGCACGCCGAGGCCAACGACTTCCGCAG 933

Qy     311 cgagtaga 318
           |||  | |
Db     934 CGACGACA 941


RESULT  13
US-09-721-822A-10
; Sequence 10, Application US/09721822A
; Patent No. 6306606
; GENERAL INFORMATION:
;  APPLICANT: Michael J. Weber
;  APPLICANT:  Jacqueline Wyatt
;  APPLICANT:  Lex M. Cowsert
;  TITLE OF INVENTION: ANTISENSE MODULATION OF MP-1 EXPRESSION
```

```
;  FILE REFERENCE: RTS-0142
;  CURRENT APPLICATION NUMBER: US/09/721,822A
;  CURRENT FILING DATE:  2000-11-22
;  NUMBER OF SEQ ID NOS: 135
; SEQ ID NO 10
;   LENGTH: 1315
;   TYPE: DNA
;   ORGANISM: Mus musculus
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (147)...(521)
;   OTHER INFORMATION:
US-09-721-822A-10


  Query Match             1.7%;  Score 52.4;  DB 4;  Length 1315;
  Best Local Similarity   79.5%;  Pred. No. 0.0041;
  Matches   62;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy   2920 tttgagagaccaatatatacagtaaataatatgtctatatgtaaaaaaaaaaaaaaaaa 2979
          |||   | ||   | |||| |  ||||| ||| | || | ||||||||||||||||||||
Db   1221 ttttttatacatgtgtataaatataaatagtatttttaaaggtaaaaaaaaaaaaaaaaa 1280

Qy   2980 aaaaaaaaaaaaaaaaa 2997
          |||||||||||||||||
Db   1281 aaaaaaaaaaaaaaaaa 1298


RESULT  14
5457037-2
;Patent No. 5457037
;     APPLICANT: TOGNONI, ANGELO;CARRERA, PAOLO;CAMERINI, BARBARA;
;GALLI, GIULIANO;LUCCHESE, GIUSEPPE;GRANDI, GUIDO;DI GENNARO, CARLO
;     TITLE OF INVENTION: CLONING OF THE GENE CODING THE ISOAMYLASE
; ENZYME AND ITS USE IN THE PRODUCTION OF SAID ENZYME
;     NUMBER OF SEQUENCES: 7
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/1,797
;       FILING DATE: 08-JAN-1993
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: 749,621
;       FILING DATE: 19-AUG-1991
;       APPLICATION NUMBER: 224,114
;       FILING DATE: 25-JUL-1988
;SEQ ID NO:2:
;       LENGTH: 2253
5457037-2


  Query Match             1.7%;  Score 51.6;  DB 6;  Length 2253;
  Best Local Similarity   49.2%;  Pred. No. 0.0079;
  Matches  314;  Conservative    0;  Mismatches  279;  Indels   45;  Gaps    5;

Qy    707 ctcaaaaggacctggtaatatatgagatgcacttgcgtggattcacgaagcatga---tt 763
          | || |||||    || || || ||| ||||  |||| || |||||  |||| ||    |
Db    524 cgcagaaggatgatgtgatctacgaggtgcatgtgcgcggcttcaccgagcaggacacct 583

Qy    764 caagcaatgtagaacatccgggtactttcattggagctgtgtcgaagcttgactatttga 823
          | | |  ||   |  |||  || || |     || || | |   |||      ||  |
Db    584 ctatccctgcgcagtatcgcggcacctattacggtgcagggctcaaggccagttacctcg 643

Qy    824 aggagcttggagttaattgtattgaattaatgccctgccatgagttcaacgagctggagt 883
               || || || |      | |||||  ||||    || ||     |  |    | |
Db    644 ccagcctgggcgtgaccgcggtggaattcctgccggtgcaggaaacgcagaatgatgcga 703

Qy    884 ac---------tcaacctcttcttccaagatgaacttttggggatattctaccataaact 934
          ||         | |  ||   | ||||   |||||  ||||| ||    |||   ||||
Db    704 acgatgtggttcccaattcagatgccaaccagaactactggggctacatgaccgagaact 763

Qy    935 tcttttcaccaatgacaagatacacatcaggcgggataaaaaactgtgggcgtgatgcca 994
           ||| || ||        | ||  | |           || ||    |   | | | |||
Db    764 acttctcgccggatcgccgctatgcct---------acaacaaggcggctggcggtccca 814

Qy    995 taaatgagttcaaaacttttgtaagagaggctcacaaacggggaattgaggtgatcctgg 1054
          ||||||  |  |  | ||    | |  ||||||    || ||  ||||   | |||
Db    815 cggcggagttccaggcgatggtgcaggcgtttcacaacgcaggcatcaaggtctacatgg 874

Qy   1055 atgttgtcttcaaccatacagctgaggg----------------------taatgagaatg 1093
          || | |||| |||||| || || || ||                     |      | |
Db    875 atatggtctacaaccacaccgccgaaggcggcacctggaccagcagtgatcccaccacgg 934

Qy   1094 gtccaatattatcatttaaggggtcgataatactacatactatatgcttgcacc---ca 1150
            | || |  || |      ||  | || ||||| || |||||   |||  |  |   ||
Db    935 ccaccatttattcgtggcgcggcttggacaataccacgtactacgagctgacctcgggca 994

Qy   1151 agggagagttttataactattctggctgtgggaataccttcaactgtaatcatcctgtgg 1210
          |   | | || ||  || |  | |||  |||    | ||||||     |  |  | ||||
Db    995 accaatacttctacgacaacacgggcattggcgcgaacttcaatacgtacaacacggtgg 1054

Qy   1211 ttcgtcaattcattgtagattgtttaagatactgggtgacggaaatgcatgttgatggtt 1270
            |   |  | || || || |   |    || |||| ||     |||   || ||||| |
Db   1055 cgcagaaccttatcgtcgactcgctggcgtattgggcgaacacgatgggcgtggatggct 1114

Qy   1271 ttcgttttgatcttgcatccataatgaccagaggttcc 1308
          |||| || || ||||| ||| |  |   ||   ||| |
Db   1115 ttcgcttcgaccttgcttccgtgctcggcaacagttgc 1152


RESULT  15
US-09-056-556-171
; Sequence 171, Application US/09056556
; Patent No. 6350456
;  GENERAL INFORMATION:
;    APPLICANT:  Reed, Steven G.
;    APPLICANT:  Skeiky, Yasir A.W.
;    APPLICANT:  Dillon, Davin C.
;    TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR THE PREVENTION AND          TI
;    NUMBER OF SEQUENCES:  241
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/056,556
;      FILING DATE:  07-APR-1998
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maki, David J.
;      REGISTRATION NUMBER:  31,392
;      REFERENCE/DOCKET NUMBER:  210121.457
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  171:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  535 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-056-556-171


  Query Match             1.7%;  Score 50.8;  DB 4;  Length 535;
  Best Local Similarity   52.2%;  Pred. No. 0.0065;
  Matches  142;  Conservative    0;  Mismatches  122;  Indels    8;  Gaps    1;

Qy      9 ccggcgccgcgcctgcgacggtggcgacccaatgcgacggcggggaaggggtcggcgag 68
          |||||||| |    |  || | |    ||    ||| ||||||||  || || | | | |
```

```
Db      2 CCGGCGCCACCGGCGGCACCGGGTTCGCCGGTGGCGCCGGCGGGGCCGGCGGGCAGGGCG 61

Qy     69 gtgtgcgccgcggttgtcgaggcggcgacgaaggtagaggacgaggggaggaggacgag 128
          || |  || | |   |  |   |  || |   ||| | |   | ||    || |||| ||
Db     62 GTATCAGCGGTGCCGGCGGCACCAACGGCTCTGGTGGCGCTGGCGGCACCGGCGGACAAG 121

Qy    129 ccggtggcggaggacaggtacgcgctcgg--------cggcgcgtgcagggtgctcgccg 180
           ||| | ||| ||    |   | ||  |         |  | |  |||  |    |||||
Db    122 GCGGCGCCGGGGGCGCTGGCGGGGCCGGCGCCGATAACCCCACCGGCATCGGCGGCGCCG 181

Qy    181 gaatgcccgcgccgctgggcgccaccgcgctcgccggcggggtcaatttcgccgtctatt 240
          |     |||        ||||   | ||    |||||||||| |   |  |||| ||  |
Db    182 GCGGCACCGGCGGCACCGGCGGAGCGGCCGGAGCCGGCGGGGCCGGTGGCGCCATCGGTA 241

Qy    241 ccggcggagccaccgccgcggcgctctgcctc 272
          |||||||  ||  || ||||| |  | || ||
Db    242 CCGGCGGCACCGGCGGCGCGGTGGGCAGCGTC 273


Search completed: July 31, 2002, 13:03:22
Job time: 17056 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: July 31, 2002, 08:17:56 ; Search time 2239.04 Seconds
(without alignments)
18065.945 Million cell updates/sec

Title: US-09-674-817A-1
Perfect score: 2997
Sequence: 1 ggtcggggccggcgccgcgc.....aaaaaaaaaaaaaaaaaaaa 2997

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		525.4	17.5	578	10	BE493792	BE493792 WHE1275_E
2	c	396	13.2	644	9	AI737641	AI737641 605036B10
3		390.2	13.0	913	10	BG444934	BG444934 GA__Ea002
4	c	390	13.0	435	10	BE422551	BE422551 WHE0055_F
5		328	10.9	469	10	BI075394	BI075394 IP1_20_A0
6	c	319	10.6	574	9	AI901664	AI901664 618008G03
7		305.6	10.2	592	10	BI934640	BI934640 EST554529
8		303.2	10.1	580	10	BI934559	BI934559 EST554448
9		278.6	9.3	558	9	AW832588	AW832588 sm14b07.y
10		271.8	9.1	547	10	BM178903	BM178903 saj60h07.
11	c	256	8.5	508	10	BE453995	BE453995 946048H01
12		250.6	8.4	759	10	BI176864	BI176864 EST517809
13	c	238.4	8.0	556	9	AW231361	AW231361 687063C04
14	c	227.6	7.6	472	10	BE475829	BE475829 946048H01
15		202.4	6.8	456	10	BI120843	BI120843 F023P78Y
16	c	200.4	6.7	704	10	BJ171788	BJ171788 BJ171788
17		192.4	6.4	383	10	T18655	T18655 5C01A06-T7
18		188.4	6.3	360	10	T15291	T15291 5c05e06-t7
19		183.4	6.1	550	9	AU036847	AU036847 AU036847
20		168.8	5.6	478	10	BF595518	BF595518 su66b03.y
21		141.8	4.7	710	10	BI933413	BI933413 EST553314
22		136	4.5	546	9	AI489006	AI489006 EST247345
23		134.6	4.5	711	10	BI923941	BI923941 EST543830
24		131.6	4.4	562	10	BF269721	BF269721 GA__Eb000
25		128	4.3	737	10	BG595915	BG595915 EST494593
26	c	126.4	4.2	627	10	BM436788	BM436788 VVA010A03
27		123.4	4.1	602	10	BE492683	BE492683 WHE0561_B
28		120	4.0	206	10	BI121740	BI121740 F045P75Y
29		108	3.6	257	10	BE525738	BE525738 M63G12STM
30	c	107.2	3.6	282	9	AV537818	AV537818 AV537818
31		106	3.5	770	10	BG887514	BG887514 EST513365
32		103.8	3.5	680	9	BE055860	BE055860 GA__Ea000
33		101.2	3.4	515	10	BG412260	BG412260 OV2_39_G0
34		95.6	3.2	638	10	BG595281	BG595281 EST493959
35		95.4	3.2	531	9	AI730342	AI730342 BNLGH1668
36		92.6	3.1	378	9	AV630278	AV630278 AV630278
37		90	3.0	568	9	AW734185	AW734185 sk80g07.y
38	c	88.2	2.9	468	12	AQ870269	AQ870269 nbeb0037D
39	c	87.6	2.9	561	12	B28627	B28627 T14K19TF TA
40	c	87.6	2.9	660	12	CNS00XRI	AL095244 Arabidops
41		87.2	2.9	514	9	AV638042	AV638042 AV638042
42		82.6	2.8	604	10	BI265380	BI265380 NF095F07I
43		76	2.5	365	9	AW697292	AW697292 NF117C04S
44	c	72.8	2.4	925	12	CNS0091P	AL053013 Drosophil
45	c	72.4	2.4	519	9	AV564985	AV564985 AV564985

ALIGNMENTS

RESULT 1
BE493792
LOCUS BE493792 578 bp mRNA linear EST 02-AUG-2000
DEFINITION WHE1275_E05_I09ZS Secale cereale anther cDNA library Secale cereale cDNA clone WHE1275_E05_I09, mRNA sequence.
ACCESSION BE493792
VERSION BE493792.1 GI:9660385
KEYWORDS EST.
SOURCE rye.
ORGANISM Secale cereale
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae ; Triticeae; Secale.
REFERENCE 1 (bases 1 to 578)
AUTHORS Anderson,O.D., Butler,E., Chao,S., Choi,D.W., Close,T.J., Fenton ,R.D., Gustafson,J.P., Han,P.S., Hsia,C.C., Kang,Y., Lazo,G.R., Miller,R., Rausch,C.J., Ross,K., Seaton,C.L. and Tong,J.C.
TITLE The structure and function of the expressed portion of the wheat genomes - Anther cDNA library from rye
JOURNAL Unpublished (2000)
COMMENT Contact: Olin Anderson
US Department of Agriculture, Agriculture Research Service, Pacific West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105595773
Fax: 5105595818
Email: oandersn@pw.usda.gov
Sequence have been trimmed to remove vector sequence and low quality sequence with phred score less than 20
Seq primer: Stratagene SK primer.
FEATURES Location/Qualifiers
source 1. .578
/organism="Secale cereale"
/cultivar="Blanco"
/db_xref="taxon:4550"
/clone="WHE1275_E05_I09"
/clone_lib="Secale cereale anther cDNA library"
/tissue_type="Anther"
/dev_stage="Adult plant before anthesis"
/lab_host="E. coli SOLR"

/note="Vector: Lambda Uni-ZAP XR, excised phagemid; Site_1: EcoRI; Site_2: XhoI; Plants were grown in the greenhouse. Anthers were harvested and pooled from early meiosis to late meiosis. The tissue, total RNA, and poly(A) RNA were prepared (Butler, Ross and Gustafson) at University of Missouri, Columbia. A cDNA library was made, and the cDNA clones were in vivo excised to give pBluescript phagemids in the TJ Close lab (Choi, Close, Fenton) at the University of California, Riverside. Plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab (all other authors)."

BASE COUNT 165 a 126 c 152 g 135 t
ORIGIN

Query Match 17.5%; Score 525.4; DB 10; Length 578;
Best Local Similarity 95.3%; Pred. No. 2.5e-49;
Matches 552; Conservative 0; Mismatches 26; Indels 1; Gaps 1;

```
Qy   2251 gggagaacaacagagatggagaaaatcacaatcttagctggaattgtggggaggaaggag 2310
          | | | |||||||||||||||||||| |||||||||||||||| ||||||||||||||||
Db      1 GAGGGGACAACAGAGATGGAGAAAAT-ACAATCTTAGCTGGAACTGTGGGGAGGAAGGAG 59

Qy   2311 aattcgcaagattgtctgtcaaaagattgaggaagaggcagatgcgcaatttctttgttt 2370
          |||||||||||||||| |||||||||||||| ||||||||||||||||||||||||||||
Db     60 AATTCGCAAGATTGTCAGTCAAAAGATTGAGAAAGAGGCAGATGCGCAATTTCTTTGTTT 119

Qy   2371 gtctcatggtttctcaaggagttccaatgttctacatgggtgatgaatatggccacacaa 2430
          |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db    120 GTCTCATGGTTTCTCAAGGAGTTCCAATGTTCTACATGGGCGATGAATATGGCCACACAA 179

Qy   2431 aaggggggcaacaacaatacatactgccatgattcttatgtcaattatttttcgctgggata 2490
          |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db    180 AAGGGGGCAACAACAATACATACTGTCATGATTCTTATGTCAATTATTTTCGCTGGGATA 239

Qy   2491 aaaaagaacaatactctgagttgcaccgattctgctgcctcatgaccaaattccgcaagg 2550
          |||||||||||||||||| ||| ||||||||||||||||||||||||||||||||||||
Db    240 AAAAAGAACAATACTCTGACTTGTACCGATTCTGCTGCCTCATGACCAAATTCCGCAAGG 299

Qy   2551 agtgcgagggtcttggccttgaggactttccaacggccaaacggctgcagtggcatggtc 2610
          |||||||||||||||||||||||||||||||||||||  ||||| ||||||||||||||||
Db    300 AGTGCGAGGGTCTTGGCCTTGAGGACTTTCCAACGGCTGAACGGTTGCAGTGGCATGGTC 359

Qy   2611 atcagcctgggaagcctgattggtctgagaatagccgattcgttgccttttccatgaaag 2670
          |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db    360 ATCAGCCTGGGAAGCCTGATTGGTCTGAGAAAAGCCGATTCGTTGCCTTTTCCATGAAAG 419

Qy   2671 atgaaagacagggcgagatctatgtggccttcaacaccagccacttaccggccgttgttg 2730
          |||||| | | || ||||||||||||||||||||||||||||||||||||||||||||||
Db    420 ATGAAACAAAAGGTGAGATCTATGTGGCCTTCAACACCAGCCACTTACCGGCCGTTGTTG 479

Qy   2731 agctcccagagcgcgcagggcgccggtgggaaccggtggtggacacaggcaagccagcac 2790
          | |||||||||||| |||||| ||||||||||||||||||||||||| |||  |||||
Db    480 AACTCCCAGAGCGCACAGGGCACCGGTGGGAACCGGTGGTGGACACAGGAAAGGAAGCAC 539

Qy   2791 catacgacttcctcaccgacgacttacctgatcgcgctc 2829
          |||| |||||||||||| ||||||||||||||||||||
Db    540 CATATGACTTCCTCACCCACGACTTACCTGATCGCGCTC 578
```

RESULT 2
AI737641/c
LOCUS AI737641 644 bp mRNA linear EST 02-FEB-2000
DEFINITION 605036B10.x2 605 - Endosperm cDNA library from Schmidt lab Zea mays cDNA, mRNA sequence.
ACCESSION AI737641
VERSION AI737641.1 GI:5069676
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.

REFERENCE 1 (bases 1 to 644)
AUTHORS Walbot,V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford University
JOURNAL Unpublished (1999)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 605036 row: B column: 10.
FEATURES Location/Qualifiers
source 1. .644
/organism="Zea mays"
/cultivar="Ohio43"
/db_xref="taxon:4577"
/clone_lib="605 - Endosperm cDNA library from Schmidt lab"
/tissue_type="nucellar, embryo, and endosperm"
/dev_stage="10-14 days post-pollination"
/lab_host="DH5(alpha)"
/note="Organ: Kernel; Vector: pAD-GAL4-2'; Site_1: EcoRI; Site_2: XhoI; Kernel endosperm cDNA library from Schmidt lab"
BASE COUNT 165 a 162 c 161 g 156 t
ORIGIN

Query Match 13.2%; Score 396; DB 9; Length 644;
Best Local Similarity 82.6%; Pred. No. 4.8e-35;
Matches 466; Conservative 0; Mismatches 95; Indels 3; Gaps 1;

```
Qy   2378 ggtttctcaaggagttccaatgttctacatgggtgatgaatatggccacacaaaaggggg 2437
          ||||||||| ||||||||||||||||||||||| |||||||||| |||||||| || ||
Db    644 GGTTTCTCAGGGAGTTCCAATGTTCTACATGGGCGATGAATATGGTCACACAAAGGGAGG 585

Qy   2438 caacaacaatacatactgccatgattcttatgtcaattattttcgctgggataaaa---a 2494
          |||||||||| ||||||||||   |||||||||||||| || |||||||| |   |
Db    584 GAACAACAATACGTACTGCCATGACCATTATGTCAATTATTTCCGTTGGGATAAGAAGGA 525

Qy   2495 agaacaatactctgagttgcaccgattctgctgcctcatgaccaaattccgcaaggagtg 2554
          |||||||| |||||| ||| || |||||||| | |||||||| |||||||||| || ||
Db    524 AGAACAATCCTCTGATTTGTACAGATTCTGCCGTCTCATGACCGAATTCCGCAAAGAATG 465

Qy   2555 cgagggtcttggccttgaggactttccaacggccaaacggctgcagtggcatggtcatca 2614
          ||   |||||||||||||||||| || ||  |  ||||| || | ||||| ||||||||
Db    464 TGAATCTCTTGGCCTTGAGGACTTCCCGACTTCAGAACGGTTGAAATGGCACGGTCATCA 405

Qy   2615 gcctgggaagcctgattggtctgagaatagccgattcgttgccttttccatgaaagatga 2674
          ||| |||||||||| ||||| |||   ||||||||||||||||  ||||||| || ||
Db    404 GCCCGGGAAGCCTGACTGGTCAGAGGCAAGCCGATTCGTTGCCTTCACCATGAAGGACGA 345

Qy   2675 aagacagggcgagatctatgtggccttcaacaccagccacttaccggccgttgttgagct 2734
          ||   | |||||||||| ||||||||||||||||| ||| | |||| ||||||| |||
Db    344 AACCAAAGGCGAGATCTACGTGGCCTTCAACACCAGTCACCTTCCGGTGGTTGTTGGGCT 285

Qy   2735 cccagagcgcgcagggcgccggtgggaaccggtggtggacacaggcaagccagcaccata 2794
          |||||||| | |||  ||| ||||| |||||||||||||| ||||||   ||||||||
Db    284 TCCAGAGCGCTCTGGGTTCCGATGGGAGCCGGTGGTGGACACCGGCAAGGAGGCACCATA 225

Qy   2795 cgacttcctcaccgacgacttacctgatcgcgctctcaccatacaccagttctcgcattt 2854
          |||||||||||||| | | | || ||||| ||| ||||| |  |||||||||| |||||
Db    224 TGACTTCCTCACCGATGGCCTGCCAGATCGTGCTGTCACCGTCTACCAGTTCTCTCATTT 165

Qy   2855 cctctactccaacctctaccccatgctcagctactcatcggtcatcctagtattgcgccc 2914
          |||| ||||||| ||||| || |||||||||||||| ||  ||||||| ||||||||||
Db    164 CCTCAACTCCAATCTCTATCCTATGCTCAGCTACTCCTCCATCATCCTTGTATTGCGCCC 105

Qy   2915 tgatgtttgagagaccaatatata 2938
          |||||| ||| |||   | ||| |
Db    104 TGATGTCTGAAAGAAGCAGATACA 81
```

```
RESULT   3
BG444934
LOCUS       BG444934                913 bp    mRNA    linear   EST 15-MAR-2001
DEFINITION  GA__Ea0026B23f Gossypium arboreum 7-10 dpa fiber library Gossypium
            arboreum cDNA clone GA__Ea0026B23f, mRNA sequence.
ACCESSION   BG444934
VERSION     BG444934.1  GI:13354586
KEYWORDS    EST.
SOURCE      Gossypium arboreum.
  ORGANISM  Gossypium arboreum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Malvales; Malvaceae; Gossypium.
REFERENCE   1  (bases 1 to 913)
  AUTHORS   Wing,R.A., Frisch,D., Yu,Y., Main,D., Rambo,T., Simmons,J., Henry
            ,D., Wood,T.C., Leslie,A. and Wilkins,T.A.
  TITLE     An integrated analysis of the genetics, development, and evolution
            of the cotton fiber
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Wing RA
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Tel: 864 656 7288
            Fax: 864 656 4293
            Email: rwing@clemson.edu
            Seq primer: TAATACGACTCACTATAGGG
            High quality sequence stop: 777.
FEATURES             Location/Qualifiers
     source          1. .913
                     /organism="Gossypium arboreum"
                     /strain="AKA"
                     /cultivar="8400"
                     /db_xref="taxon:29729"
                     /clone="GA__Ea0026B23f"
                     /clone_lib="Gossypium arboreum 7-10 dpa fiber library"
                     /tissue_type="Fibers isolated from bolls harvested 7-10
                     dpa"
                     /lab_host="E. coli"
                     /note="Vector: pBK-CMV; Site_1: EcoRI; Site_2: XhoI"
BASE COUNT      254 a    158 c    235 g    266 t
ORIGIN
```

Query Match 13.0%; Score 390.2; DB 10; Length 913;
Best Local Similarity 72.2%; Pred. No. 1.9e-34;
Matches 542; Conservative 0; Mismatches 193; Indels 16; Gaps 2;

```
Qy    586 ataagccgaggggagtatggtgttccagcgcgtggtaacaattgctggcctcagatggct 645
          |||||| |||||||||| ||| | |  || || ||   | ||||| ||||| || |||||
Db     26 ATAAGCAGAGGGGAGTTTGGGGCTTTAGGGCCTGAAGATAATTGTTGGCCCCAAATGGCC 85

Qy    646 ggcatgatccctcttccatatagcacgtttgattgggaaggcgacctacctctaagatat 705
          || ||| | |||  | || |      ||||||||||||||| ||  ||||| | ||| ||
Db     86 GGAATGGTGCCTACTTCAGAATATCAGTTTGATTGGGAAGGTGATTTACCTTTGAGACAT 145

Qy    706 cctcaaaaggacctggtaatatatgagatgcacttgcgtggattcacgaagcatgattca 765
          ||  | |  || ||  |||| || || |||||  |||||||||  || |||||||| ||
Db    146 CCAGACAGAGATCTCATAATTTACGAAATGCATGTGCGTGGATATACAAAGCATGAATCT 205

Qy    766 agcaatgtagaacatccgggtactttcattggagctgtgtcgaagcttgactatttgaag 825
          ||       | ||  ||| || || |    ||| |  |||   |||||||||| | || |||
Db    206 AGTGGGACAAAATTTCCTGGCACGTATCGTGGTGTGGTGGAAAAGCTTGACCACTTAAAG 265

Qy    826 gagcttggagttaattgtattgaattaatgccctgccatgagttcaacgagctggagtac 885
          || |||||||| || || || |||||||||| || || || ||||| ||| |||||||||
Db    266 GAACTTGGAGTCAACTGCATAGAATTAATGCCATGTCAAGAATTCAATGAGTTGGAGTAC 325

Qy    886 tcaacctcttcttc--------------caagatgaacttttggggatattctaccata 930
          |  | ||    |||                ||| ||||||||||||||||||| || ||
Db    326 TACAGCTACAATTCTGTTTTGGGTGACTATAAGTTGAACTTTTGGGGATATTCAACAATC 385

Qy    931 aacttctttttcaccaatgacaagatacacatcaggcgggataaaaaaactgtgggcgtgat 990
          || |  ||||| || |||| ||| ||  ||||    || ||    | |||||| ||||||
Db    386 AATTATTTTTCCCCCATGATAAGGTATTCATCTTCTGGTATTCGTAGCTGTGGTCGTGAT 445

Qy    991 gccataaatgagttcaaaacttttgtaagagaggctcacaaacggggaattgaggtgatc 1050
          || || ||||| ||||||  | |||| | ||| || || ||||| ||||||||||| ||
Db    446 GCAATTAATGAATTCAAATATCTTGTTAAAGAAGCACATAAACGCGGAATTGAGGTTATA 505

Qy   1051 ctggatgttgtcttcaaccatacagctgagggtaatgagaatggtccaatattatcattt 1110
          |||||||||| ||||| || |||||||| ||  ||||||| ||||  |  || ||||||
Db    506 ATGGATGTTGTTTTCAATCACACAGCTGAAGGGCATGAGAAAGGTCTCAGTTTGTCATTT 565

Qy   1111 aaggggtcgataatactacatactatatgcttgcacccaagggagagttttataactat 1170
          |  || || |||||   |   || || ||||| ||||| |||||||||| || || |||
Db    566 AGAGGTGTTGATAACTGTGTCTATTACATGCTGGCACCTAAGGGAGAGTACTACAATTAT 625

Qy   1171 tctggctgtgggaataccttcaactgtaatcatcctgtggttcgtcaattcattgtagat 1230
          || || |||||||| || ||||||||| | |||||||||||| |||||||| ||  ||||
Db    626 TCAGGGTGTGGGAACACATTCAACTGTGACCATCCTGTGGTGCGTCAATTTATATTAGAC 685

Qy   1231 tgtttaagatactgggtgacggaaatgcatgttgatggttttcgttttgatcttgcatcc 1290
          ||  ||||||| ||||| || |||||||||| ||||| | ||| ||||||||||| |||
Db    686 TGCCTAAGATATTGGGTAACAGAAATGCATGTCGATGGGTGTCGCTTTGATCTTGCTTCC 745

Qy   1291 ataatgaccagaggttccagtctgtgggatc 1321
          || ||||||||  ||  | |||| |||||||
Db    746 AT-ATGACCAGGAGTAGCGGTCTTTGGGATC 775
```

```
RESULT   4
BE422551/c
LOCUS       BE422551                435 bp    mRNA    linear   EST 24-JUL-2000
DEFINITION  WHE0055_F11_K21ZS Wheat endosperm cDNA library Triticum aestivum
            cDNA clone WHE0055_F11_K21, mRNA sequence.
ACCESSION   BE422551
VERSION     BE422551.1  GI:9420310
KEYWORDS    EST.
SOURCE      bread wheat.
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
            ; Triticeae; Triticum.
REFERENCE   1  (bases 1 to 435)
  AUTHORS   Altenbach,S., Anderson,O.D., Chao,S., Galili,G., Han,P.S., Hsia
            ,C.C., Kang,Y., Lazo,G.R., Miller,R., Rausch,C.J., Seaton,C.L. and
            Tong,J.C.
  TITLE     The structure and function of the expressed portion of the wheat
            genomes - Endosperm cDNA library
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Olin Anderson
            US Department of Agriculture, Agriculture Research Service, Pacific
            West Area, Western Regional Research Center
            800 Buchanan Street, Albany, CA 94710, USA
            Tel: 5105595773
            Fax: 5105595818
            Email: oandersn@pw.usda.gov
            Sequence have been trimmed to remove vector sequence and low
            quality sequence with phred score less than 20
            Seq primer: Stratagene SK primer.
FEATURES             Location/Qualifiers
     source          1. .435
                     /organism="Triticum aestivum"
                     /cultivar="Cheyenne"
                     /db_xref="taxon:4565"
                     /clone="WHE0055_F11_K21"
                     /clone_lib="Wheat endosperm cDNA library"
                     /tissue_type="Endosperm"
                     /dev_stage="5 to 30 days post anthesis seed"
                     /lab_host="E. coli SOLR"
                     /note="Vector: Lambda ZAP II, excised phagemid; Site_1:
                     EcoRI; Seeds collected, endosperm isolated, and RNA
                     prepared by Susan  Altenbach. Library constructed by
                     Stratagene, Inc. Plasmid DNA  preparations and DNA
```

```
                    sequencing were performed in the OD Anderson lab."
BASE COUNT      108 a    101 c    124 g    102 t
ORIGIN


  Query Match             13.0%;  Score 390;  DB 10;  Length 435;
  Best Local Similarity   100.0%;  Pred. No. 2.5e-34;
  Matches  390;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   2573 ggactttccaacggccaaacggctgcagtggcatggtcatcagcctgggaagcctgattg 2632
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    435 GGACTTTCCAACGGCCAAACGGCTGCAGTGGCATGGTCATCAGCCTGGGAAGCCTGATTG 376

Qy   2633 gtctgagaatagccgattcgttgccttttccatgaaagatgaaagacagggcgagatcta 2692
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    375 GTCTGAGAATAGCCGATTCGTTGCCTTTTCCATGAAAGATGAAAGACAGGGCGAGATCTA 316

Qy   2693 tgtggccttcaacaccagccacttaccggccgttgttgagctcccagagcgcgcagggcg 2752
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    315 TGTGGCCTTCAACACCAGCCACTTACCGGCCGTTGTTGAGCTCCCAGAGCGCGCAGGGCG 256

Qy   2753 ccggtgggaaccggtggtggacacaggcaagccagcaccatacgacttcctcaccgacga 2812
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    255 CCGGTGGGAACCGGTGGTGGACACAGGCAAGCCAGCACCATACGACTTCCTCACCGACGA 196

Qy   2813 cttacctgatcgcgctctcaccatacaccagttctcgcatttcctctactccaacctcta 2872
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    195 CTTACCTGATCGCGCTCTCACCATACACCAGTTCTCGCATTTCCTCTACTCCAACCTCTA 136

Qy   2873 ccccatgctcagctactcatcggtcatcctagtattgcgccctgatgtttgagagaccaa 2932
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    135 CCCCATGCTCAGCTACTCATCGGTCATCCTAGTATTGCGCCCTGATGTTTGAGAGACCAA 76

Qy   2933 tatatacagtaaataatatgtctatatgta 2962
          |||||||||||||||||||||||||||||
Db     75 TATATACAGTAAATAATATGTCTATATGTA 46


RESULT   5
BI075394
LOCUS       BI075394                469 bp    mRNA    linear   EST 20-JUN-2001
DEFINITION  IP1_20_A07.b1_A002 Immature pannicle 1 (IP1) Sorghum bicolor cDNA,
            mRNA sequence.
ACCESSION   BI075394
VERSION     BI075394.1  GI:14514051
KEYWORDS    EST.
SOURCE      sorghum.
  ORGANISM  Sorghum bicolor
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE   1  (bases 1 to 469)
  AUTHORS   Klein,R.R., Cordonnier-Pratt,M.-M., Gingle,A., Sudman,M. and Pratt
            ,L.H.
  TITLE     An EST database from Sorghum: developing preanthesis pannicles
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Cordonnier-Pratt MM
            Department of Botany
            The University of Georgia
            Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
            Tel: 706 542 1860
            Fax: 706 542 1805
            Email: mmpratt@uga.edu
            Sequences have been trimmed to exclude PolyA, vector and regions
            below Phred quality 16. The threshold for high quality sequence is
            20. Three-prime sequences, which are obtained with PolyTMix or T7
            sequencing primer, are presented as the reverse complement.
            Seq primer: JEN REV
            High quality sequence stop: 463
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .469
                     /organism="Sorghum bicolor"
                     /cultivar="BTx623"
                     /db_xref="taxon:4558"
                     /clone_lib="Immature pannicle 1 (IP1)"
                     /note="Organ: Developing preanthesis pannicles; Vector:
                     pBluescript II SK(-) from Lambda Zap II; Site_1: XhoI;
                     Site_2: EcoRI; The library was made from poly-A RNA in the
                     cloning vector lambda ZAP II.  Clones to be sequenced were
                     prepared by mass excision."
BASE COUNT      129 a     99 c    111 g    130 t
ORIGIN


  Query Match             10.9%;  Score 328;  DB 10;  Length 469;
  Best Local Similarity   82.5%;  Pred. No. 1.7e-27;
  Matches  376;  Conservative    0;  Mismatches   80;  Indels    0;  Gaps    0;

Qy    570 ttatgctaaggcagtgataagccgaggggagtatggtgttccagcgcgtggtaacaattg 629
          || |  | |||||||| |||||||||  | |||||||| || |  | |||| | |||||
Db     14 TTTTCTTTAGGCAGTGGTAAGCCGAGGTAAATATGGTGTGCCTGGACCTGGTGATAATTG 73

Qy    630 ctggcctcagatggctggcatgatccctcttccatatagcacgtttgattgggaaggcga 689
          |||||||||  |||||||| ||||||||||||||  |||| |||||| | ||| |||||||
Db     74 CTGGCCTCAAATGGCTGGTATGATCCCTCTTCCGCATAGTACGTTTAACTGGCAAGGCGA 133

Qy    690 cctacctctaagatatcctcaaaaggacctggtaatatatgagatgcacttgcgtggatt 749
          ||||||  || |  |||| ||| |||||||| || || ||||| ||||| |||||||| ||
Db    134 CCTACCCCTGAAGTATCATCAGAAGGACCTTGTCATCTATGAAATGCATTTGCGTGGGTT 193

Qy    750 cacgaagcatgattcaagcaatgtagaacatccgggtactttcattggagctgtgtcgaa 809
          ||| |||||||||||||||| | | |||| || || |||| |||||| |||||||| ||
Db    194 CACCAAGCATGATTCAAGCAACGCAAAACACCCAGGAACTTACATTGGTGCTGTGTCAAA 253

Qy    810 gcttgactatttgaaggagcttggagttaattgtattgaattaatgccctgccatgagtt 869
          |||||||  || | |||||  |||||||| |||||||| || ||||||||||||||||||
Db    254 GCTTGACCATCTAAAGGAACTTGGAGTCAATTGTATAGAGTTAATGCCCTGCCATGAGTT 313

Qy    870 caacgagctggagtactcaacctcttcttccaagatgaacttttggggatattctaccat 929
          |||||||||  |||||||  | ||| |||||||||||||||| |||||  |||||  || ||
Db    314 CAACGAGCTAGAGTACTTCAGCTCCTCTTCCAAGATGAACTTCTGGGGGTATTCCACAAT 373

Qy    930 aaacttcttttcaccaatgacaagatacacatcaggcgggataaaaaactgtgggcgtga 989
          ||| || || |||||||||  |||||||  | ||| | || |||| | ||| ||| || |
Db    374 AAATTTTTTCTCACCAATGGCAAGATATTCGTCAAGTGGCATAAGAGACTCTGGACGCGG 433

Qy    990 tgccataaatgagttcaaaacttttgtaagagaggc 1025
          |||||||||||| || ||||||||||||||  |||||
Db    434 TGCCATAAATGAATTTAAAACTTTTGTAAGGGAGGC 469


RESULT   6
AI901664/c
LOCUS       AI901664                574 bp    mRNA    linear   EST 27-JUL-1999
DEFINITION  618008G03.x1 618 - Inbred Tassel cDNA Library Zea mays cDNA, mRNA
            sequence.
ACCESSION   AI901664
VERSION     AI901664.1  GI:5607997
KEYWORDS    EST.
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 574)
  AUTHORS   Walbot,V.
  TITLE     Maize ESTs from various cDNA libraries sequenced at Stanford
            University
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
```

```
            Email: walbot@stanford.edu
            Plate: 618008  row: G  column: 03.
FEATURES              Location/Qualifiers
     source           1. .574
                      /organism="Zea mays"
                      /cultivar="Ohio43"
                      /db_xref="taxon:4577"
                      /clone_lib="618 - Inbred Tassel cDNA Library"
                      /tissue_type="tassel"
                      /dev_stage="tassel length from 0.1 to 2.5 cm"
                      /lab_host="XLOLR"
                      /note="Organ: tassel; Vector: pAD-GAL4-2.1 (Hybrizap);
                      Inbred tassel library from Schmidt lab"
BASE COUNT      159 a    127 c    147 g    140 t      1 others
ORIGIN


  Query Match             10.6%;  Score 319;  DB 9;  Length 574;
  Best Local Similarity   79.0%;  Pred. No. 1.6e-26;
  Matches  392;  Conservative    0;  Mismatches  101;  Indels    3;  Gaps    1;

Qy   2446 atacatactgccatgattcttatgtcaattatttttcgctgggataaaa---aagaacaat 2502
           |||| ||||||||||||   ||||||||||||||| || |||||||| |   |||||||||
Db    574 ATACGTACTGCCATGACCATTATGTCAATTATTTCCGTTGGGATAAGAAGGAAGAACAAT 515

Qy   2503 actctgagttgcaccgattctgctgcctcatgaccaaattccgcaaggagtgcgagggtc 2562
            || ||| ||| || |||||||| | ||||||||||||||||||||||| || ||   ||
Db    514 CCTNTGATTTGTACAGATTCTGCCGTCTCATGACCAAATTCCGCAAGGAATGTGAATCTC 455

Qy   2563 ttggccttgaggactttccaacggccaaacggctgcagtggcatggtcatcagcctggga 2622
           |||||||||||||| || || ||  |  ||||| |||  | ||| | |||||||||| ||||
Db    454 TTGGCCTTGAGGATTTCCCGACTTCAGAACGGTTGCGATAGCACGATCATCAGCCCGGGA 395

Qy   2623 agcctgattggtctgagaatagccgattcgttgccttttccatgaaagatgaaagacagg 2682
           ||||||| ||||| |||   ||||||||||||||||  ||||||| || ||||   | |
Db    394 AGCCTGACTGGTCAGAGGCAAGCCGATTCGTTGCCTTCACCATGAAGGACGAAACCAAAG 335

Qy   2683 gcgagatctatgtggccttcaacaccagccacttaccggccgttgttgagctcccagagc 2742
             ||||||||  |||||||||||||||||| ||| | ||||  || || | |   | |||||
Db    334 ATGAGATCTACGTGGCCTTCAACACCAGTCACCTGCCGGTGGTCGTCGGGTGTCTAGAGC 275

Qy   2743 gcgcagggcgccggtgggaaccggtggtggacacaggcaagccagcaccatacgacttcc 2802
           || |  ||  ||| ||||| |||||||||||||||| ||||||  |||||||| |||||||
Db    274 GCTCTAGGTTCCGATGGGAGCCGGTGGTGGACACCGGCAAGGAGGCACCATATGACTTCC 215

Qy   2803 tcaccgacgacttacctgatcgcgctctcaccatacaccagttctcgcatttcctctact 2862
           ||||||| | | | || ||||| ||| ||||| |  |||||||||| ||||| ||| |||
Db    214 TCACCGATGGCCTGCCAGATCGTGCTGTCACCGTCTACCAGTTCTCTCATTTTCTCAACT 155

Qy   2863 ccaacctctaccccatgctcagctactcatcggtcatcctagtattgcgccctgatgttt 2922
           |||| ||||| || |||||||||||||| ||  ||||||| |||||||||||||||| |
Db    154 CCAATCTCTATCCTATGCTCAGCTACTCCTCCATCATCCTTGTATTGCGCCCTGATGTCT 95

Qy   2923 gagagaccaatatata 2938
           || |||     ||| |
Db     94 GAAAGAAATGGATACA 79


RESULT   7
BI934640
LOCUS       BI934640                 592 bp    mRNA    linear   EST 18-OCT-2001
DEFINITION  EST554529 tomato flower, anthesis Lycopersicon esculentum cDNA
            clone cTOD2006 5' end, mRNA sequence.
ACCESSION   BI934640
VERSION     BI934640.1  GI:16249112
KEYWORDS    EST.
SOURCE      tomato.
  ORGANISM  Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
            Lycopersicon.
REFERENCE   1  (bases 1 to 592)
  AUTHORS   van der Hoeven,R.S., Bezzerides,J.L., Karamycheva,S.A., Tsai,J.,
            Utterback,T., Van Aken,S., Ronning,C.M., Nierman,W., Fraser,C.M.,
            Martin,G.B., Giovannoni,J.J. and Tanksley,S.D.
  TITLE     Generation of ESTs from tomato flower tissue, anthesis (2001)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: CUGI
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Email: http://www.genome.clemson.edu/orders/index.html
            This clone is available through the Clemson University Genomics
            Institute
            Seq primer: T3.
FEATURES              Location/Qualifiers
     source           1. .592
                      /organism="Lycopersicon esculentum"
                      /cultivar="TA496"
                      /db_xref="taxon:4081"
                      /clone="cTOD2006"
                      /clone_lib="tomato flower, anthesis"
                      /tissue_type="flower"
                      /dev_stage="anthesis"
                      /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
                      XhoI; supplier: Cornell University; sequencing: The
                      Institute for Genomic Research; Flower buds and flowers
                      were taken from greenhouse plants (4-8 wks old, TA496).
                      They were immediately frozen in liquid nitrogen and then
                      size-separated while remaining frozen."
BASE COUNT      171 a    101 c    129 g    191 t
ORIGIN


  Query Match             10.2%;  Score 305.6;  DB 10;  Length 592;
  Best Local Similarity   72.4%;  Pred. No. 4.7e-25;
  Matches  417;  Conservative    0;  Mismatches  144;  Indels   15;  Gaps    1;

Qy    672 gtttgattgggaaggcgacctacctctaagatatcctcaaaaggacctggtaatatatga 731
          |||||||||||||| || ||||| || |  | | |||| |   | ||||||| |||||
Db     16 GTTTGATTGGGAAGGAGATCTACCACTGAAGTTTTCTCAGAGAAATCTGGTAATCTATGA 75

Qy    732 gatgcacttgcgtggattcacgaagcatgattcaagcaatgtagaaacatccgggtacttt 791
          |||||  | |||||||| || || ||||| || ||  |   | || |||| |||||||
Db     76 AATGCATGTTCGTGGATTTACAAATCATGAGTCGAGTGAAACAAAATATCCTGGTACTTA 135

Qy    792 cattggagctgtgtcgaagcttgactatttgaaggagcttggagttaattgtattgaatt 851
          | |||| | ||||  ||| ||||| | |||||||| ||||| || || ||||| ||  |
Db    136 CCTTGGTGTTGTGGAGAAACTTGATCACTTGAAGGAACTTGGTGTCAACTGTATAGAGCT 195

Qy    852 aatgccctgccatgagttcaacgagctggagtactcaacctcttcttc---------- 899
          ||||||||| || |||||||| ||||||||||||  |  | |   ||
Db    196 AATGCCCTGTCACGAGTTCAATGAGCTGGAGTACTATAGTTATAACTCTGTATTGGGCGA 255

Qy    900 ---caagatgaacttttggggatattctaccataaacttcttttcaccaatgacaagata 956
              |||| | |||||||||||| |||||||  | || |||||||| ||||||  ||||||
Db    256 CTACAAGTTTAACTTTTGGGGCTATTCTACTGTCAATTTCTTTTCTCCAATGGGAAGATA 315

Qy    957 cacatcaggcgggataaaaaactgtgggcgtgatgccataaatgagttcaaaacttttgt 1016
          | |||| |  ||  |||  || ||||| |  | ||| ||||| || || ||   | ||||
Db    316 CTCATCTGCTGGTCTAAGTAATTGTGGCCTCGGTGCAATAAACGAATTTAAGTATCTTGT 375

Qy   1017 aagagaggctcacaaacggggaattgaggtgatcctggatgttgtcttcaaccatacagc 1076
          |  || || || ||||| ||||| ||||| ||| |||||||||| ||||| || || ||
Db    376 CAAGGAAGCACATAAACGTGGAATCGAGGTTATCATGGATGTTGTTTTCAATCACACTGC 435

Qy   1077 tgagggtaatgagaatggtccaatattatcatttaaggggtcgataatactacatacta 1136
          ||| || ||||| |||||||| ||| ||||||||  ||  | || || | |   |  ||
Db    436 TGAAGGAAATGAAAATGGTCCCATACTATCATTTAGAGGCATTGACAACAGTGTGTTTTA 495

Qy   1137 tatgcttgcacccaagggagagttttataactattctggctgtgggaataccttcaactg 1196
          || ||| || || ||||| ||  |||   |||| || || ||||| ||||||||||||||
Db    496 TACGCTAGCTCCTAAGGGTGAAATTTTACACTACTCAGGATGTGGAAATACCTTCAACTG 555

Qy   1197 taatcatcctgtggttcgtcaattcattgtagattg 1232
```

```
          |||| ||||  | || ||||||||| || || |||||
Db    556 TAATAATCCCATTGTACGTCAATTTATAGTGGATTG 591


RESULT   8
BI934559
LOCUS       BI934559                 580 bp    mRNA    linear   EST 18-OCT-2001
DEFINITION  EST554448 tomato flower, anthesis Lycopersicon esculentum cDNA
            clone cTOD2005 5' end, mRNA sequence.
ACCESSION   BI934559
VERSION     BI934559.1  GI:16249031
KEYWORDS    EST.
SOURCE      tomato.
  ORGANISM  Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
            Lycopersicon.
REFERENCE   1  (bases 1 to 580)
  AUTHORS   van der Hoeven,R.S., Bezzerides,J.L., Karamycheva,S.A., Tsai,J.,
            Utterback,T., Van Aken,S., Ronning,C.M., Nierman,W., Fraser,C.M.,
            Martin,G.B., Giovannoni,J.J. and Tanksley,S.D.
  TITLE     Generation of ESTs from tomato flower tissue, anthesis (2001)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: CUGI
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Email: http://www.genome.clemson.edu/orders/index.html
            This clone is available through the Clemson University Genomics
            Institute
            Seq primer: T3.
FEATURES             Location/Qualifiers
     source          1. .580
                     /organism="Lycopersicon esculentum"
                     /cultivar="TA496"
                     /db_xref="taxon:4081"
                     /clone="cTOD2005"
                     /clone_lib="tomato flower, anthesis"
                     /tissue_type="flower"
                     /dev_stage="anthesis"
                     /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
                     XhoI; supplier: Cornell University; sequencing: The
                     Institute for Genomic Research; Flower buds and flowers
                     were taken from greenhouse plants (4-8 wks old, TA496).
                     They were immediately frozen in liquid nitrogen and then
                     size-separated while remaining frozen."
BASE COUNT      168 a    100 c    125 g    187 t
ORIGIN


  Query Match             10.1%;  Score 303.2;  DB 10;  Length 580;
  Best Local Similarity   72.9%;  Pred. No. 8.7e-25;
  Matches  411;  Conservative    0;  Mismatches  138;  Indels   15;  Gaps    1;

Qy    672 gtttgattgggaaggcgacctacctctaagatatcctcaaaaggacctggtaatatatga 731
          |||||||||||||||| || ||||| || |  | | |||| |   | |||||||| |||||
Db     16 GTTTGATTGGGAAGGAGATCTACCACTGAAGTTTTCTCAGAGAAATCTGGTAATCTATGA 75

Qy    732 gatgcacttgcgtggattcacgaagcatgattcaagcaatgtagaacatccgggtacttt 791
           |||||  | |||||||| || || ||||| || ||  |   | || |||| |||||||
Db     76 AATGCATGTTCGTGGATTTACAAATCATGAGTCGAGTGAAACAAAATATCCTGGTACTTA 135

Qy    792 cattggagctgtgtcgaagcttgactatttgaaggagcttggagttaattgtattgaatt 851
          | |||| | ||||  ||| ||||| | |||||||| ||||| || || ||||| ||  |
Db    136 CCTTGGTGTTGTGGAGAAACTTGATCACTTGAAGGAACTTGGTGTCAACTGTATAGAGCT 195

Qy    852 aatgccctgccatgagttcaacgagctggagtactcaacctcttcttc------------ 899
          |||||||| || |||||||| ||||||||||||||  |  | |   ||
Db    196 AATGCCCTGTCACGAGTTCAATGAGCTGGAGTACTATAGTTATAACTCTGTATTGGGCGA 255

Qy    900 ---caagatgaacttttggggatattctaccataaacttcttttcaccaatgacaagata 956
             |||| | |||||||||||| ||||||||  | || |||||||| ||||||  ||||||

Db    256 CTACAAGTTTAACTTTTGGGGCTATTCTACTGTCAATTTCTTTTCTCCAATGGGAAGATA 315

Qy    957 cacatcaggcgggataaaaaactgtgggcgtgatgccataaatgagttcaaaacttttgt 1016
          | |||| |  ||  |||  || |||||  |  | ||| ||||| || || ||   | ||||
Db    316 CTCATCTGCTGGTCTAAGTAATTGTGGCCTCGGTGCAATAAACGAATTTAAGTATCTTGT 375

Qy   1017 aagagaggctcacaaacggggaattgaggtgatcctggatgttgtcttcaaccatacagc 1076
            |  || || || ||||| ||||| ||||| ||| |||||||||| ||||| || || ||
Db    376 CAAGGAAGCACATAAACGTGGAATCGAGGTTATCATGGATGTTGTTTTCAATCACACTGC 435

Qy   1077 tgagggtaatgagaatggtccaatattatcatttaaggggtcgataatactacatacta 1136
          ||| || ||||| |||||||| ||| |||||||||  ||  | || || | |   |  ||
Db    436 TGAAGGAAATGAAAATGGTCCCATACTATCATTTAGAGGCATTGACAACAGTGTGTTTTA 495

Qy   1137 tatgcttgcacccaagggagagttttataactattctggctgtgggaataccttcaactg 1196
          || ||| || || ||||| || ||||| ||||| || || ||||| ||||||||||||||
Db    496 TACGCTAGCTCCTAAGGGTGAATTTTACAACTACTCAGGATGTGGAAATACCTTCAACTG 555

Qy   1197 taatcatcctgtggttcgtcaatt 1220
          |||| ||||  | || ||||||||
Db    556 TAATAATCCCATTGTACGTCAATT 579


RESULT   9
AW832588
LOCUS       AW832588                 558 bp    mRNA    linear   EST 03-DEC-2001
DEFINITION  sm14b07.y1 Gm-c1027 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
            Gm-c1027-8846 5' similar to TR:O22637 O22637 SU1 ISOAMYLASE. ;,
            mRNA sequence.
ACCESSION   AW832588
VERSION     AW832588.1  GI:7926562
KEYWORDS    EST.
SOURCE      soybean.
  ORGANISM  Glycine max
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
            Glycine.
REFERENCE   1  (bases 1 to 558)
  AUTHORS   Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna
            ,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
            Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
            ,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
            ,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
            ,R., Waterston,R. and Wilson,R.
  TITLE     Public Soybean EST Project
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Shoemaker R/Public Soybean EST Project
            Public Soybean EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available through: ResGen, Invitrogen Corp. 2130
            South Memorial Parkway Huntsville, AL 35801 For further information
            call: (800)-533-4363 or contact via email: ccu@resgen.com
            High quality sequence stop: 411.
FEATURES             Location/Qualifiers
     source          1. .558
                     /organism="Glycine max"
                     /db_xref="taxon:3847"
                     /clone="GENOME SYSTEMS CLONE ID: Gm-c1027-8846"
                     /clone_lib="Gm-c1027"
                     /tissue_type="cotyledons of 3- and 7-day-old Williams
                     seedlings"
                     /lab_host="DH10B"
                     /note="Vector: pBluescript II SK+; Site_1: EcoRI; Site_2:
                     XhoI; This cDNA library was constructed from mRNA isolated
                     from cotyledons of 3- and 7-day-old Williams seedlings
                     which were propagated on paper towels with distilled
                     water.  The cotyledons were flash-frozen in liquid
                     nitrogen, then lyophilized for 72 hours.  Unequal amounts
```

of mRNA was used for cDNA synthesis. Stratagene's cDNA Synthexix Kit (catalog number 200401) was used to synthesize the cDNA. First- stranded synthesis was performed with 5-methyl dCTP, hence the ligated cDNA was hemimethylated. A modification of Stratagene's first-strand synthesis primer was used. An anchor nucleotide (V=A, C, or G) was added to the 3' end of the primer [GAGAGAGAGAGAGAGAGAGAACTAGTCTCGAG(T)18] to anchor the primer at the 5' end of the poly(A) tract. After second- strand synthesis, the cDNA ends were filled in with cloned Pfu DNA, ligated to EcoRI adapters and subsequently phosphorylated. The XhoI site within the first-strand synthesis primer was then restricted by digestion with XhoI; all XhoI sites in the cDNA would be protected by their hemimethylated status. The cDNA constructs were size-fractionated with a 500 bp cutoff, using GibcoBRL Life Technologies' cDNA Size Fractionation column. The column eluent was then ligated into Stratagene's pBluescript(tm) II XR Predigested vector (pBluescript II SK(+) that has been digested with EcoRI and XhoI, and phosphorylated by Stratagene). 97% of the white and blue colonies appear to contain recombinant plasmids with cDNA inserts, based on size (n=30). This library was constructed by Dr. Paul Keim and Dr. Virginia Coryell."

BASE COUNT 156 a 127 c 111 g 164 t
ORIGIN

Query Match 9.3%; Score 278.6; DB 9; Length 558;
Best Local Similarity 70.0%; Pred. No. 4.6e-22;
Matches 390; Conservative 0; Mismatches 164; Indels 3; Gaps 1;

```
Qy   2349 cagatgcgcaatttctttgtttgtctcatggtttctcaaggagttccaatgttctacatg 2408
          || ||||| ||||| ||| ||| |||||||||||| || ||||||||||| | || |||
Db      1 CAAATGCGGAATTTTTTTCTTTCTCTCATGGTTTCCCAGGGAGTTCCAATGATATATATG 60

Qy   2409 ggtgatgaatatggccacacaaaaggggcaacaacaatacatactgccatgattcttat 2468
          || |||||||||| |||||||||| || || |||||||| || || || |||  ||||
Db     61 GGCGATGAATATGGACACACAAAAGGAGGAAATAACAATACCTATTGTCACGATAATTAT 120

Qy   2469 gtcaattattttcgctgggataaaaaagaacaatac---tctgagttgcaccgattctgc 2525
          ||||| || |  ||||| ||||| ||| ||| |   || || ||   | |||| ||
Db    121 CATAATTACTTCCAATGGGACAAAAAGGAAGAATCCTCATCAGACTTCTTCAGATTTTGT 180

Qy   2526 tgcctcatgaccaaattccgcaaggagtgcgagggtcttggccttgaggactttccaacg 2585
          |||| ||||| || |||||| |||| || ||    || ||| | |  ||||| |||||
Db    181 CGCCTTATGACTAAGTTCCGCCAGGAATGTGAATCGCTAGGCTTAGCTGACTTCCCAACC 240

Qy   2586 gccaaacggctgcagtggcatggtcatcagcctgggaagcctgattggtctgagaatagc 2645
          |  |  ||||||||||||||||||||   ||||| ||||| || |||||||| |  |||
Db    241 TCTGAGAGGCTGCAGTGGCATGGTCATTTTCCTGGAAAGCCAGACTGGTCTGAAACCAGC 300

Qy   2646 cgattcgttgccttttccatgaaagatgaaagacagggcgagatctatgtggccttcaac 2705
          || || || || | | |||||  ||||  |    |||| || || ||  | || |||||
Db    301 CGTTTTGTGGCTTGTACCATGGTAGATTCAGTGAAGGGAGAAATATACATTGCTTTCAAT 360

Qy   2706 accagccacttaccggccgttgttgagctcccagagcgcgcagggcgccggtgggaaccg 2765
          |  || || |||||  |  |||  | || |||||  | ||  |  |||||||||
Db    361 ATGAGTCATTTACCTTTCACAGTTACCTTGCCGGAGCGTCCTGGATACAAATGGGAACCT 420

Qy   2766 gtggtggacacaggcaagccagcaccatacgacttcctcaccgacgacttacctgatcgc 2825
          | || |||||  |||||||  ||||||| || ||||||||    ||| | ||||   |
Db    421 CTTGTAGACACCAGCAAGCCTACACCATATGATTTCCTCACTCCTGACCTTCCTGGAAGA 480

Qy   2826 gctctcaccatacaccagttctcgcatttcctctactccaacctctaccccatgctcagc 2885
          | | |  ||||||| ||||   | || || ||  || |||  | || |||||||| ||
Db    481 GATATTGCCATACAACAGTATGCTCAGTTTCTGGACGCCAATATGTATCCCATGCTTAGA 540

Qy   2886 tactcatcggtcatcct 2902
          || || ||  | |||||
Db    541 TATTCTTCCATTATCCT 557
```

RESULT 10
BM178903
LOCUS BM178903 547 bp mRNA linear EST 06-DEC-2001
DEFINITION saj60h07.y1 Gm-c1072 Glycine max cDNA clone SOYBEAN CLONE ID: Gm-c1072-4213 5' similar to TR:O22637 O22637 SU1 ISOAMYLASE. ;, mRNA sequence.
ACCESSION BM178903
VERSION BM178903.1 GI:17402121
KEYWORDS EST.
SOURCE soybean.
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
REFERENCE 1 (bases 1 to 547)
AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.
TITLE Public Soybean EST Project
JOURNAL Unpublished (1999)
COMMENT Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL 35801 For further information call: (800)-533-4363 or contact: ccu@resgen.com web site: www.resgen.com
Seq primer: -40RP from Gibco
High quality sequence stop: 422.
FEATURES Location/Qualifiers
source 1. .547
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="SOYBEAN CLONE ID: Gm-c1072-4213"
/clone_lib="Gm-c1072"
/tissue_type="seedlings induced for symptoms of SDS (Sudden Death Syndrome) disease"
/dev_stage="2-3 weeks old"
/lab_host="DH10B"
/note="Vector: pBluescript II SK+; Site_1: EcoRI; Site_2: XhoI; The cDNA library wms constructed from mRNA isolated from 2-3 week old seedlings that were induced for symptoms of SDS (Sudden Death Syndrome) disease by the translocation of culture filtrate of Fusarium solani f. sp. glycines (Plant Cell Report 18:375-380). Cultivar PI 567374 is partially resistant to the disease SDS. Plant tissue (expanded leaves, folded leaves, and new shoots) were collected at 1, 6, 24, and 48 hrs. after inoculation and their mRNA pooled equally for cDNA construction. The library was prepared using the Stratagene pBluescript II SK(+) library construction kit. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with an XhoI restriction site. EcoRI adaptors were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA insert is protected from XhoI digestion via methylation during first strand synthesis. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into E.coli ElectroMax DH10B host cells. Plants were inoculated by Shuxian Li (Glen Hartman lab, University of Illinois). Library was constructed by Steve Clough (Lila Vodkin lab, University of Illinois)."
BASE COUNT 151 a 124 c 111 g 161 t

```
ORIGIN


  Query Match             9.1%;  Score 271.8;  DB 10;  Length 547;
  Best Local Similarity   69.8%;  Pred. No. 2.6e-21;
  Matches  382;  Conservative    0;  Mismatches  162;  Indels    3;  Gaps    1;

Qy   2354 gcgcaatttctttgtttgtctcatggtttctcaaggagttccaatgttctacatgggtga 2413
          ||| ||||| ||| ||| ||||||||||||| || ||||||||||||| | || ||||| ||
Db      1 GCGGAATTTTTTTCTTTCTCTCATGGTTTCCCAGGGAGTTCCAATGATATATATGGGCGA 60

Qy   2414 tgaatatggccacacaaaaggggcaacaacaatacatactgccatgattcttatgtcaa 2473
          ||||||||| |||||||||| || || |||||||| || || || ||| |||| ||
Db     61 TGAATATGGACACACAAAAGGAGGAAATAACAATACCTATTGTCACGATAATTATCATAA 120

Qy   2474 ttattttcgctgggataaaaaagaacaatac---tctgagttgcaccgattctgctgcct 2530
          ||| || |  ||||| ||||| ||| ||| |   || || ||   | |||| ||  ||||
Db    121 TTACTTCCAATGGGACAAAAAGGAAGAATCCTCATCAGACTTCTTCAGATTTTGTCGCCT 180

Qy   2531 catgaccaaattccgcaaggagtgcgagggtcttggccttgaggactttccaacggccaa 2590
          ||||| || |||||| |||| || ||    || ||| | |  ||||| |||||  |  |
Db    181 TATGACTAAGTTCCGCCAGGAATGTGAATCGCTAGGCTTAGCTGACTTCCCAACCTCTGA 240

Qy   2591 acggctgcagtggcatggtcatcagcctgggaagcctgattggtctgagaatagccgatt 2650
          ||||||||||||||||||||    ||||| ||||| || |||||||| |  ||||| ||
Db    241 GAGGCTGCAGTGGCATGGTCATTTTCCTGGAAAGCCAGACTGGTCTGAAACCAGCCGTTT 300

Qy   2651 cgttgccttttccatgaaagatgaaagacagggcgagatctatgtggccttcaacaccag 2710
          || || | | |||||  ||||  |    |||| || || ||  | || |||||  |  ||
Db    301 TGTGGCTTGTACCATGGTAGATTCAGTGAAGGGAGAAATATACATTGCTTTCAATATGAG 360

Qy   2711 ccacttaccggccgttgttgagctcccagagcgcgcagggcgccggtgggaaccggtggt 2770
          || |||||   |  |||    | || |||||  | ||   |  ||||||||  | ||
Db    361 TCATTTACCTTTCACAGTTACCTTGCCGGAGCGTCCTGGATACAAATGGGAACCTCTTGT 420

Qy   2771 ggacacaggcaagccagcaccatacgacttcctcaccgacgacttacctgatcgcgctct 2830
          |||||  |||||||  ||||||| || ||||||||    ||| | ||||   | | | |
Db    421 AGACACCAGCAAGCCTACACCATATGATTTCCTCACTCCTGACCTTCCTGGAAGAGATAT 480

Qy   2831 caccatacaccagttctcgcatttcctctactccaacctctaccccatgctcagctactc 2890
          |||||||  ||||   | || || ||  || ||||  | || ||||||||  || || ||
Db    481 TGCCATACAACAGTATGCTCAGTTTCTGGACGCCAATATGTATCCCATGCTTAGTTATTC 540

Qy   2891 atcggtc 2897
          ||  ||
Db    541 TTCCATC 547


RESULT  11
BE453995/c
LOCUS       BE453995                508 bp    mRNA    linear   EST 26-JUL-2000
DEFINITION  946048H01.x1 946 - tassel primordium prepared by Schmidt lab Zea
            mays cDNA, mRNA sequence.
ACCESSION   BE453995
VERSION     BE453995.1  GI:9461969
KEYWORDS    EST.
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 508)
  AUTHORS   Walbot,V.
  TITLE     Maize ESTs from various cDNA libraries sequenced at Stanford
            University
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Plate: 946048  row: H  column: 01.
FEATURES             Location/Qualifiers
     source          1. .508
                     /organism="Zea mays"
                     /cultivar="OH43"
                     /db_xref="taxon:4577"
                     /clone_lib="946 - tassel primordium prepared by Schmidt
                     lab"
                     /tissue_type="tassels"
                     /dev_stage="just after the transition from vegetative to
                     inflorescence development"
                     /lab_host="XLOLR"
                     /note="Organ: tassels; Vector: HybriZAP; Site_1: EcoRI;
                     Site_2: XhoI; George Chuck dissected immature tassels
                     between 1mm and 3mm.  Sharon  Stanfield prepared the cDNA
                     library in HybriZAP.  Sample insert size  range was 350 bp
                     to 3 Kb with a 1 Kb average."
BASE COUNT      134 a    116 c    137 g    121 t
ORIGIN


  Query Match             8.5%;  Score 256;  DB 10;  Length 508;
  Best Local Similarity   78.3%;  Pred. No. 1.5e-19;
  Matches  307;  Conservative    0;  Mismatches   85;  Indels    0;  Gaps    0;

Qy   2547 aaggagtgcgagggtcttggccttgaggactttccaacggccaaacggctgcagtggcat 2606
          ||||| || ||   ||||||||||||||| ||| || ||  |  ||||| |||  | |||
Db     508 AAGGAATGTGAATCTCTTGGCCTTGAGGCCTTCCCGACTTCAGAACGGTTGCGATAGCAC 449

Qy   2607 ggtcatcagcctgggaagcctgattggtctgagaatagccgattcgttgccttttccatg 2666
          | ||||||||| |||||||||| ||||| |||   |||||||||||||||||  |||||
Db     448 GATCATCAGCCCGGGAAGCCTGACTGGTCAGAGGCAAGCCGATTCGTTGCCTTCACCATG 389

Qy   2667 aaagatgaaagacagggcgagatctatgtggccttcaacaccagccacttaccggccgtt 2726
          || || ||||   | |||||||||||| |||||||||||||||| ||| |  |||  ||
Db     388 AAGGACGAAACCAAAGGCGAGATCTACGTGGCCTTCAACACCAGTCACCTGTCGGTGGTC 329

Qy   2727 gttgagctcccagagcgcgcagggcgccggtgggaaccggtggtggacacaggcaagcca 2786
          || | |||   ||||||| | |||  ||| ||||| ||||||||||||||| ||||||
Db     328 GTCGGGCTTTTAGAGCGCTCTGGGTTCCGATGGGAGCCGGTGGTGGACACCGGCAAGGAG 269

Qy   2787 gcaccatacgacttcctcaccgacgacttacctgatcgcgctctcaccatacaccagttc 2846
          |||||||| |||||||||| || | | | || ||||| ||| ||||| |  |||||||
Db     268 GCACCATATGACTTCCTCACTGATGGCCTGCCAGATCGTGCTGTCACCGTCTACCAGTTC 209

Qy   2847 tcgcatttcctctactccaacctctaccccatgctcagctactcatcggtcatcctagta 2906
          || |||||||| ||||||| ||||| || ||||||| |||||| ||  ||||||| |||
Db     208 TCTCATTTCCTCAACTCCAATCTCTATCCTATGCTCAACTACTCCTCCATCATCCTTGTA 149

Qy   2907 ttgcgccctgatgtttgagagaccaatatata 2938
          ||||||||||||| ||| |||     ||| |
Db     148 TTGCGCCCTGATGTCTGAAAGAAAGGGATACA 117


RESULT  12
BI176864
LOCUS       BI176864                759 bp    mRNA    linear   EST 09-JUL-2001
DEFINITION  EST517809 cSTE Solanum tuberosum cDNA clone cSTE6P8 5' sequence,
            mRNA sequence.
ACCESSION   BI176864
VERSION     BI176864.1  GI:14642675
KEYWORDS    EST.
SOURCE      potato.
  ORGANISM  Solanum tuberosum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
REFERENCE   1  (bases 1 to 759)
  AUTHORS   van der Hoeven,R., Bezzerides,J., Bachem,C., Visser,R., Cho,J.,
            Chiemingo,A., Bougri,O., Buell,C.R., Ronning,C., Tanksley,S. and
            Baker,B.
  TITLE     Generation of ESTs from in vitro grown microtubers
  JOURNAL   Unpublished (2001)
```

```
COMMENT     Contact: Cathy Ronning
            The Institute for Genomic Research
            For clone info: please contact Research Genetics, Libraries
            Division tel 1-800-711-6195, email cdna@resgen.com
            Seq primer: M13F-R.
FEATURES             Location/Qualifiers
     source          1. .759
                     /organism="Solanum tuberosum"
                     /cultivar="Bintje"
                     /db_xref="taxon:4113"
                     /clone="cSTE6P8"
                     /clone_lib="cSTE"
                     /tissue_type="axillary buds of stem explants; growing
                     sink-tubers"
                     /dev_stage="7, 8 and 10 days"
                     /lab_host="SOLR"
                     /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
                     XhoI; Tissue supplied by Christian Bachem and Richard
                     Visser (Department of Plant Breeding, Wageningen
                     University, The Netherlands). The cSTA libraries will
                     attempt to capture the induction and initiation/initial
                     growth of the tuber in an in vitro system as described in
                     Bachem et al. (Plant Journal, 1996). Small microtubers
                     develop from axillary buds attached to stem explants when
                     placed on a high sucrose medium (10%). Visible
                     morphological changes occur synchronously at day five in
                     the axillary buds. The first library, cSTA (1-20) consists
                     of axillary buds harvested on days 1-3. This targets
                     those genes involved in induction of the microtubers. The
                     following libraries, cSTA (21-40) and cSTA (41-60),
                     capture genes involved in tuber initiation and outgrowth.
                     This library is noted as P3 in Tanksley lab notebooks."
BASE COUNT      190 a    128 c    210 g    231 t
ORIGIN
```

Query Match 8.4%; Score 250.6; DB 10; Length 759;
Best Local Similarity 68.5%; Pred. No. 5.2e-19;
Matches 346; Conservative 0; Mismatches 159; Indels 0; Gaps 0;

```
Qy    395 aggatagggtgaccgaggaggttccccttgacccccctgatgaatcggaccgggaacgtgt 454
          || | || |||||||||| |  ||  ||| || || ||    |||   || ||  | || |
Db    216 AGAAGAGAGTGACCGAGCAAATTTTCCTGGATCCTCTAGCTAATAAAACTGGAGATGTAT 275

Qy    455 ggcatgtcttcatcgaaggcgagctgcacaacatgctttacgggtacaggttcgacggca 514
          ||||||| ||| |  | || ||  |  | || |||||  || ||  ||||  || || || |
Db    276 GGCATGTGTTCCTTAAGGGAGATTTTGAGAATATGCTATATGGCTACAAATTTGATGGGA 335

Qy    515 cctttgctcctcactgcgggcactaccttgatgtttccaatgtcgtggtggatccttatg 574
            ||   |||| |    || |||||| ||||   |||  |  | ||| ||||||||||||
Db    336 AATTCTGTCCTGAAGAAGGACACTACTTTGACTCTTCGCAGATAGTGTTGGATCCTTATG 395

Qy    575 ctaaggcagtgataagccgaggggagtatggtgttccagcgcgtggtaacaattgctggc 634
          | |||||  |  ||||| |||| || |||||||||   || ||  |   |  |||| ||||
Db    396 CCAAGGCTATAGTAAGCAGAGGAGAATATGGTGTATTAGGGCCAGAGGATGATTGTTGGC 455

Qy    635 ctcagatggctggcatgatccctcttccatatagcacgtttgattgggaaggcgacctac 694
          | |  ||||||||||||| | |||  | | | |     ||||||||||||||| || ||||
Db    456 CCCCAATGGCTGGCATGGTACCTTCTGCTTCTGATCAGTTTGATTGGGAAGGAGATCTAC 515

Qy    695 ctctaagatatcctcaaaaggacctggtaatatatgagatgcacttgcgtggattcacga 754
          | || |  | ||| || |  || || ||||| ||||| |||||  | ||||| || || |
Db    516 CACTGAAGTTTCCACAGAGAGATCTTGTAATCTATGAAATGCATGTTCGTGGGTTTACTA 575

Qy    755 agcatgattcaagcaatgtagaacatccgggtactttcattggagctgtgtcgaagcttg 814
          | ||||| || ||  |   | || |||| ||||||| | |||| | ||||  ||| ||||
Db    576 ATCATGAGTCGAGTGAAACAAAATATCCTGGTACTTACCTTGGTGGTGTGGAGAAACTTG 635

Qy    815 actatttgaaggagcttggagttaattgtattgaattaatgccctgccatgagttcaacg 874
          |  | |||||||| ||||| || || ||||| ||  |||||||||| || ||||||||  |
Db    636 ATCACTTGAAGGAACTTGGTGTCAACTGTATAGAGCTAATGCCCTGTCACGAGTTCAATG 695

Qy    875 agctggagtactcaacctcttcttc 899
```

```
          ||||||||||||  |  | |   ||
Db    696 AGCTGGAGTACTATAGTTATAACTC 720
```

```
RESULT  13
AW231361/c
LOCUS       AW231361                556 bp    mRNA    linear   EST 30-MAR-2000
DEFINITION  687063C04.x1 687 - Early embryo from Delaware Zea mays cDNA, mRNA
            sequence.
ACCESSION   AW231361
VERSION     AW231361.1  GI:6563829
KEYWORDS    EST.
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 556)
  AUTHORS   Walbot,V.
  TITLE     Maize ESTs from various cDNA libraries sequenced at Stanford
            University
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Plate: 687063  row: C  column: 04.
FEATURES             Location/Qualifiers
     source          1. .556
                     /organism="Zea mays"
                     /cultivar="Illinois High Oil"
                     /db_xref="taxon:4577"
                     /clone_lib="687 - Early embryo from Delaware"
                     /tissue_type="embryo"
                     /dev_stage="14, 21, 28, and 35 days after pollination"
                     /lab_host="E. coli SOLR"
                     /note="Organ: embryo; Vector: pBluescript SK; Site_1: XhoI
                     ; Site_2: EcoRI; Library was prepared by Statagene using
                     the Uni-ZAP XR system (Stratagene BN937328-12). Clones
                     were picked by a Q-bot after blue/white selection
                     (ampicillin resistance - use 100 micrograms/microliter).
                     Developed from a pool of equal amounts of RNA from
                     developing embryos sampled at 14, 21, 28 and 35 days after
                     pollination of the Illinois High Oil Maize Strain Cycle
                     90. This closed strain has been selected for high oil
                     concentration for 90 generations and originates from the
                     1890s era open pollinated variety Burr's White"
BASE COUNT      143 a    157 c    147 g    108 t      1 others
ORIGIN
```

Query Match 8.0%; Score 238.4; DB 9; Length 556;
Best Local Similarity 79.6%; Pred. No. 1.3e-17;
Matches 281; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

```
Qy   2576 ctttccaacggccaaacggctgcagtggcatggtcatcagcctgggaagcctgattggtc 2635
          ||| || ||  |  ||||| || | ||||| |||||||||| |||||||||| |||||
Db    556 CTTCCCGACTTCAGAACGGTTGAAATGGCACGGTCATCAGCCCGGGAAGCCTGACTGGTC 497

Qy   2636 tgagaatagccgattcgttgccttttccatgaaagatgaaagacagggcgagatctatgt 2695
           |||   |||||||||||||||||  |||||||  |  |||    | ||||||||||  ||
Db    496 AGAGGCAAGCCGATTCGTTGCCTTCGCCATGAAGGCCGAAGCCAAAGGCGAGATCTGCGT 437

Qy   2696 ggccttcaacaccagccacttaccggccgttgttgagctcccagagcgcgcagggcgccg 2755
          |||||||| |||||| ||| | ||||  ||||| |  ||| |||||||||  | |||  |||
Db    436 GGCCTTCAGCACCAGTCACCTTCCGGTGGTTGTCGGGCTTCCAGAGCGCTCTGGGTTCCG 377

Qy   2756 gtgggaaccggtggtggacacaggcaagccagcaccatacgacttcctcaccgacgactt 2815
          ||||| |||||||||| | | ||||||   |||||||| ||||||||||||||| | | |
Db    376 ATGGGAGCCGGTGGTGGCCGCCGGCAAGGAGGCACCATATGACTTCCTCACCGATGGCCT 317
```

```
Qy   2816 acctgatcgcgctctcaccatacaccagttctcgcatttcctctactccaacctctaccc 2875
          || ||||| ||| ||||| |  |||||||||| | ||||||| ||||||| ||||| ||
Db    316 NCCAGATCGTGCTGTCACCGTCTACCAGTTCTCTCGTTTCCTCAACTCCAATCTCTATCC 257

Qy   2876 catgctcagctactcatcggtcatcctagtattgcgccctgatgtttgagaga 2928
          |||||||||||||| ||  ||||||| |||||||||||||||||| ||| |||
Db    256 TATGCTCAGCTACTCCTCCATCATCCTTGTATTGCGCCCTGATGTCTGAAAGA 204


RESULT  14
BE475829/c
LOCUS       BE475829                472 bp    mRNA    linear   EST 28-JUL-2000
DEFINITION  946048H01.x2 946 - tassel primordium prepared by Schmidt lab Zea
            mays cDNA, mRNA sequence.
ACCESSION   BE475829
VERSION     BE475829.1  GI:9566320
KEYWORDS    EST.
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 472)
  AUTHORS   Walbot,V.
  TITLE     Maize ESTs from various cDNA libraries sequenced at Stanford
            University
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Plate: 946048  row: H  column: 01.
FEATURES             Location/Qualifiers
     source          1. .472
                     /organism="Zea mays"
                     /cultivar="OH43"
                     /db_xref="taxon:4577"
                     /clone_lib="946 - tassel primordium prepared by Schmidt
                     lab"
                     /tissue_type="tassels"
                     /dev_stage="just after the transition from vegetative to
                     inflorescence development"
                     /lab_host="XLOLR"
                     /note="Organ: tassels; Vector: HybriZAP; Site_1: EcoRI;
                     Site_2: XhoI; George Chuck dissected immature tassels
                     between 1mm and 3mm.  Sharon  Stanfield prepared the cDNA
                     library in HybriZAP.  Sample insert size  range was 350 bp
                     to 3 Kb with a 1 Kb average."
BASE COUNT      125 a    106 c    128 g    113 t
ORIGIN


  Query Match             7.6%;  Score 227.6;  DB 10;  Length 472;
  Best Local Similarity   78.6%;  Pred. No. 2.1e-16;
  Matches  272;  Conservative    0;  Mismatches   74;  Indels    0;  Gaps    0;

Qy   2593 ggctgcagtggcatggtcatcagcctgggaagcctgattggtctgagaatagccgattcg 2652
          || |||  | ||| | ||||||||| |||||||||| ||||| |||   ||||||||||
Db    472 GGTTGCGATAGCACGATCATCAGCCCGGGAAGCCTGACTGGTCAGAGGCAAGCCGATTCG 413

Qy   2653 ttgccttttccatgaaagatgaaagacagggcgagatctatgtggccttcaacaccagcc 2712
          |||||||  ||||||| || ||||   | |||||||||||| |||||||||||||||| |
Db    412 TTGCCTTCACCATGAAGGACGAAACCAAAGGCGAGATCTACGTGGCCTTCAACACCAGTC 353

Qy   2713 acttaccggccgttgttgagctcccagagcgcgcagggcgccggtgggaaccggtggtgg 2772
          || |  |||  || || | |||   ||||||| | |||  ||| ||||| ||||||||||
Db    352 ACCTGTCGGTGGTCGTCGGGCTTTTAGAGCGCTCTGGGTTCCGATGGGAGCCGGTGGTGG 293

Qy   2773 acacaggcaagccagcaccatacgacttcctcaccgacgacttacctgatcgcgctctca 2832
          |||| ||||||   |||||||| |||||||||| || | | | || ||||| ||| |||
Db    292 ACACCGGCAAGGAGGCACCATATGACTTCCTCACTGATGGCCTGCCAGATCGTGCTGTCA 233

Qy   2833 ccatacaccagttctcgcatttcctctactccaacctctaccccatgctcagctactcat 2892
          || |  |||||||||| |||||||| ||||||| ||||| || |||||||| || || |
Db    232 CCGTCTACCAGTTCTCTCATTTCCTCAACTCCAATCTCTATCCTATGCTCAGATAATCCT 173

Qy   2893 cggtcatcctagtattgcgccctgatgtttgagagaccaatatata 2938
          |  ||||||| ||||||||||||||| ||| |||     ||| |
Db    172 CCATCATCCTTGTATTGCGCCCTGATGTCTGAAAGAAAGGGATACA 127


RESULT  15
BI120843
LOCUS       BI120843                456 bp    mRNA    linear   EST 31-DEC-2001
DEFINITION  F023P78Y Populus flower cDNA library Populus balsamifera subsp.
            trichocarpa cDNA, mRNA sequence.
ACCESSION   BI120843
VERSION     BI120843.1  GI:18004818
KEYWORDS    EST.
SOURCE      Populus balsamifera subsp. trichocarpa.
  ORGANISM  Populus balsamifera subsp. trichocarpa
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Malpighiales; Salicaceae; Populus.
REFERENCE   1  (bases 1 to 456)
  AUTHORS   Hertzberg,M., Aspeborg,H., Erlandsson,R., Bjorkbacka,H., Hiltonen
            ,T., Karlsson,J., Teeri,T., Gustafsson,P., Bahlerao,R., Jansson,S.,
            Nilsson,O., Sundberg,B., Nilsson,P., Uhlen,M., Sandberg,G. and
            Lundeberg,J.
  TITLE     Gene expression in Populus
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Erlandsson R
            Department of Biotechnology
            Royal Institute of Technology
            Teknikringen 30, Stockholm S-10044, Sweden
            Tel: 46 8 790 8287
            Fax: 46 8 245452
            Email: rikerl@biochem.kth.se.
FEATURES             Location/Qualifiers
     source          1. .456
                     /organism="Populus balsamifera subsp. trichocarpa"
                     /db_xref="taxon:3694"
                     /clone_lib="Populus flower cDNA library"
                     /note="Organ: flower"
BASE COUNT      150 a     85 c     99 g    119 t      3 others
ORIGIN


  Query Match             6.8%;  Score 202.4;  DB 10;  Length 456;
  Best Local Similarity   74.1%;  Pred. No. 1.3e-13;
  Matches  335;  Conservative    0;  Mismatches  108;  Indels    9;  Gaps    6;

Qy   2140 atcaggcaggaggaaggaaaccttggcacagtatcaactttgtatgtgcacatgatggat 2199
          | |||| |||||||||||||||| ||| |||| |||||||||||||||||||||||||| |
Db      6 ACCAGGAAGGAGGAAGGAAACCATGGAACAGCATCAACTTTGTATGTGCACATGATGGTT 65

Qy   2200 ttacactggctgatttggtaacatataataagaagtacaatttaccaaatgg-ggagaac 2258
          ||||  |||||||||| || |||||||| |||||| | || ||  ||||||| | || ||
Db     66 TTACTTTGGCTGATTTAGTGACATATAACAAGAAGCATAACTTGGCAAATGGCGAAGCAC 125

Qy   2259 aacagagatggagaaa-atcacaatcttagctggaattgtgg-ggaggaaggagaattcg 2316
          ||||  |||||||||| |||| ||   ||||||||||||||   ||||    ||||| |
Db    126 AACAATGATGGAGAAACATCATAACAATAGCTGGAATTGTGGCACAGGAGTTTGAATTTG 185

Qy   2317 caagattgtctgtcaaaagattgaggaagagg--cagatgcgcaatttctttgtttgtct 2374
          | ||  | || || || |    |   ||| | |  || ||| | |||||||  | |||||
Db    186 CCAGCATTTCAGTGAAGAANNTTGCGAAAACGNACAAATGAGAAATTTCTTCCTGTGTCT 245

Qy   2375 catggtttctcaaggagttccaatgttctacatgggtgatgaatatggccacacaaaagg 2434
          ||||||||| ||||| || |||||| | ||||||||||||||||||| ||||||||||
Db    246 CATGGTTTCACAAGGTGTCCCAATGATATACATGGGTGATGAATATGGTCACACAAAAGG 305
```

```
Qy   2435 gggcaacaacaatacatactgccatgattcttatgtcaattattttcgctgggataaaaa 2494
          ||| ||||| |||||||| ||||||||| |        |     |||    |||||| ||
Db    306 GGGAAACAA-AATACATATTGCCATGATACCCTATATTAACTACTTCCGCTGGATAAGAA 364

Qy   2495 agaacaatac---tctgagttgcaccgattctgctgcctcatgaccaaattccgcaagga 2551
           ||| ||| |   || || ||     ||||  || ||||||||||||| |||||| | ||
Db    365 GGAAGAATCCTCATCAGACTTCTTTAGATTTGGCCGCCTCATGACCAAGTTCCGCCATGA 424

Qy   2552 gtgcgagggtcttggccttgaggactttccaa 2583
           || |||   | | |  |  | ||||| ||||
Db    425 ATGTGAGTCCCCTTGGTTGAATGACTTCCCAA 456
```

Search completed: July 31, 2002, 10:50:03
Job time: 9127 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: July 31, 2002, 10:50:09 ; Search time 61.77 Seconds
(without alignments)
1373.812 Million cell updates/sec

Title: US-09-674-817A-2
Perfect score: 4212
Sequence: 1 SGPAPRLRRWRPNATAGKGV.....SNLYPMLSYSSVILVLRPDV 764

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	4212	100.0	764	21	AAY50819	Wheat isoamylase p
2	3527.5	83.7	818	20	AAY17523	Zea mays SU1 starc
3	3029	71.9	793	20	AAY00869	S. tuberosum isoam
4	2919.5	69.3	931	20	AAY00872	Original S. tubero
5	2918	69.3	568	20	AAY06918	Partial hexaploid
6	2612	62.0	606	18	AAW36602	S. tuberosum debra
7	1881	44.7	429	20	AAY04382	Potato isoamylase
8	1577	37.4	766	20	AAY00870	S. tuberosum isoam
9	1305.5	31.0	666	20	AAY34991	C. pneumoniae prot
10	1206.5	28.6	618	20	AAY37184	Protein involved i
11	1204	28.6	836	22	AAG92051	C glutamicum prote
12	1187	28.2	726	20	AAW95636	Rhodothermus marin
13	1187	28.2	726	21	AAY78515	Isoamylase amino a
14	1062.5	25.2	718	22	AAU60903	Propionibacterium
15	1032.5	24.5	458	20	AAY00873	Original S. tubero
16	980.5	23.3	654	22	AAU38219	Salmonella typhi c
17	971	23.1	630	22	AAB79424	Corynebacterium gl
18	915.5	21.7	878	20	AAY00868	S. tuberosum isoam
19	910.5	21.6	942	20	AAY00871	Original S. tubero
20	889.5	21.1	3096	20	AAY06919	Hexaploid wheat DB
21	825	19.6	774	19	AAW35390	Flavobacterium odo
22	810	19.2	777	17	AAR92692	Isoamylase precurs
23	810	19.2	777	19	AAW53885	Full length Flavob
24	807	19.2	746	17	AAR92693	Isoamylase mature
25	807	19.2	746	19	AAW53886	Mature Flavobacter
26	802.5	19.1	776	21	AAY78516	Isoamylase amino a
27	781.5	18.6	750	10	AAP94419	Primary transcript
28	708.5	16.8	475	22	ABG24650	Novel human diagno
29	666	15.8	745	10	AAP90615	Polypeptide with i
30	636.5	15.1	491	22	AAB79425	Corynebacterium gl
31	552.5	13.1	915	21	AAY78514	Pullulanase amino
32	552.5	13.1	928	22	AAE05719	Bacillus deramific
33	550.5	13.1	928	22	AAE05709	Bacillus deramific
34	549.5	13.0	928	22	AAE05693	Bacillus deramific
35	547.5	13.0	928	22	AAE05704	Bacillus deramific
36	546.5	13.0	922	19	AAW37372	Trimmed enzyme pro
37	545.5	13.0	928	22	AAE05720	Bacillus deramific
38	544.5	12.9	815	22	AAE05695	Bacillus deramific
39	544.5	12.9	817	22	AAE05694	Bacillus deramific
40	544.5	12.9	928	22	AAE05690	Bacillus deramific
41	544.5	12.9	928	22	AAE05697	Bacillus deramific
42	544.5	12.9	928	22	AAE05698	Bacillus deramific
43	544.5	12.9	928	22	AAE05699	Bacillus deramific
44	544.5	12.9	928	22	AAE05700	Bacillus deramific
45	544.5	12.9	928	22	AAE05708	Bacillus deramific

ALIGNMENTS

RESULT 1
AAY50819
ID AAY50819 standard; Protein; 764 AA.
XX
AC AAY50819;
XX
DT 18-FEB-2000 (first entry)
XX
DE Wheat isoamylase protein fragment.
XX
KW Isoamylase; wheat; transgenic plant; starch production; food; baking;
KW pastry; packaging material; glucose; glucan; paper; pulp; adhesive;
KW textile building material; soil stabilizer; wetting agent; fertilizer;
KW plant-protection; cosmetic; flocculant.
XX
OS Triticum aestivum.
XX
PN DE19820608-A1.
XX
PD 11-NOV-1999.
XX
PF 08-MAY-1998; 98DE-1020608.
XX
PR 08-MAY-1998; 98DE-1020608.
XX
PA (AGRE) HOECHST-SCHERING AGREVO GMBH.
XX
PI Loerz H, Luetticke S, Abel G;
XX
DR WPI; 2000-024509/03.
DR N-PSDB; AAZ24492.
XX
PT New nucleic acid encoding isoamylase from wheat and related transgenic
PT plants producing starch with altered properties -

```
XX
PS   Claim 1a; Page 15-16; 18pp; German.
XX
CC   This invention describes a novel wheat isoamylase. Transgenic plants,
CC   specifically wheat, that contain (I) are used for production of starch,
CC   used particularly in foods, particularly baked and pastry goods and for
CC   making packaging materials or disposable items. Starch may also be used
CC   as starting materials for glucose or glucan components (e.g. for
CC   fermentation or further chemical conversion); in paper and pulp
CC   production; as adhesives; in textiles; in preparation of gypsum-based
CC   building materials; as soil stabilizer; as wetting agent etc. in
CC   fertilizer and plant-protection compositions; as binder (in
CC   pharmaceuticals, cosmetics, coal briquetting and casting sand); as
CC   flocculant in soil or coal slurries; as rubber and leather additives; and
CC   for production of synthetic polymers, e.g. polyurethane films. Transgenic
CC   plants with increased/decreased production of (I) produce starches with
CC   altered physical and/or chemical properties, e.g. amylose/amylopectin
CC   ratios; degree of branching; mean chain length; phosphate content;
CC   gelatinization properties; gel- or film-forming properties, or starch
CC   grain size or structure, e.g. easier, and less expensive, digestion by
CC   enzymes. This sequence represents the wheat isoamylase protein described
CC   in the invention.
XX
SQ   Sequence   764 AA;


  Query Match             100.0%;  Score 4212;  DB 21;  Length 764;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  764;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 SGPAPRLRRWRPNATAGKGVGEVCAAVVEAATKVEDEGEEDEPVAEDRYALGGACRVLAG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 sgpaprlrrwrpnatagkgvgevcaavveaatkvedegeedepvaedryalggacrvlag 60

Qy      61 MPAPLGATALAGGVNFAVYSGGATAAALCLFTPEDLKADRVTEEVPLDPLMNRTGNVWHV 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 mpaplgatalaggvnfavysggataaalclftpedlkadrvteevpldplmnrtgnvwhv 120

Qy     121 FIEGELHNMLYGYRFDGTFAPHCGHYLDVSNVVVDPYAKAVISRGEYGVPARGNNCWPQM 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 fiegelhnmlygyrfdgtfaphcghyldvsnvvvdpyakavisrgeygvpargnncwpqm 180

Qy     181 AGMIPLPYSTFDWEGDLPLRYPQKDLVIYEMHLRGFTKHDSSNVEHPGTFIGAVSKLDYL 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 agmiplpystfdwegdlplrypqkdlviyemhlrgftkhdssnvehpgtfigavskldyl 240

Qy     241 KELGVNCIELMPCHEFNELEYSTSSSKMNFWGYSTINFFSPMTRYTSGGIKNCGRDAINE 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 kelgvncielmpchefneleystsssksmnfwgystinffspmtrytsggiknegrdaine 300

Qy     301 FKTFVREAHKRGIEVILDVVFNHTAEGNENGPILSFKGVDNTTYYMLAPKGEFYNYSGCG 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 fktfvreahkrgievildvvfnhtaegnengpilsfkgvdnttyymlapkgefynysgcg 360

Qy     361 NTFNCNHPVVRQFIVDCLRYWVTEMHVDGFRFDLASIMTRGSSLWDPVNVYGAPIEGDMI 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 ntfncnhpvvrqfivdclrywvtemhvdgfrfdlasimtrgsslwdpvnvygapiegdmi 420

Qy     421 TTGTPLVTPPLIDMISNDPILGGVKLIAEAWDAGGLYQVGQFPHWNVWSEWNGKYRDIVR 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 ttgtplvtpplidmisndpilggvkliaeawdagglyqvgqfphwnvwsewngkyrdivr 480

Qy     481 QFIKGTDGFAGGFAECLCGSPHLYQAGGRKPWHSINFVCAHDGFTLADLVTYNKKYNLPN 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 qfikgtdgfaggfaeclcgsphlyqaggrkpwhsinfvcahdgftladlvtynkkynlpn 540

Qy     541 GENNRDGENHNLSWNCGEEGEFARLSVKRLRKRQMRNFFVCLMVSQGVPMFYMGDEYGHT 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 gennrdgenhnlswncgeegefarlsvkrlrkrqmrnffvclmvsqgvpmfymgdeyght 600

Qy     601 KGGNNNTYCHDSYVNYFRWDKKEQYSELHRFCCLMTKFRKECEGLGLEDFPTAKRLQWHG 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 kggnnntychdsyvnyfrwdkkeqyselhrfcclmtkfrkeceglgledfptakrlqwhg 660

Qy     661 HQPGKPDWSENSRFVAFSMKDERQGEIYVAFNTSHLPAVVELPERAGRRWEPVVDTGKPA 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 hqpgkpdwsensrfvafsmkderqgeiyvafntshlpavvelperagrrwepvvdtgkpa 720

Qy     721 PYDFLTDDLPDRALTIHQFSHFLYSNLYPMLSYSSVILVLRPDV 764
           |||||||||||||||||||||||||||||||||||||||||||||
Db     721 pydfltddlpdraltihqfshflysnlypmlsyssvilvlrpdv 764


RESULT   2
AAY17523
ID   AAY17523 standard; Protein; 818 AA.
XX
AC   AAY17523;
XX
DT   04-AUG-1999  (first entry)
XX
DE   Zea mays SU1 starch debranching enzyme.
XX
KW   Zea mays; SU1; starch debranching enzyme; maize; endosperm; corn;
KW   bacterial isoamylase; alpha-amylase; starch hydrolytic enzyme;
KW   sugary 1.
XX
OS   Zea mays.
XX
PN   US5912413-A.
XX
PD   15-JUN-1999.
XX
PF   24-MAR-1995;   95US-0410784.
XX
PR   24-MAR-1995;   95US-0410784.
XX
PA   (IOWA ) UNIV IOWA STATE RES FOUND INC.
XX
PI   James MG,  Myers AM;
XX
DR   WPI; 1999-357234/30.
DR   N-PSDB; AAX76375.
XX
PT   SU1 starch debranching enzyme
XX
PS   Claim 1; Fig 1; 39pp; English.
XX
CC   The present sequence represents a starch debranching enzyme active in
CC   maize endosperm (Zea mays), designated SU1. SU1 protein is useful as
CC   a replacement for the bacterial and fungal enzymes currently used in
CC   the starch processing industry.
XX
SQ   Sequence   818 AA;


  Query Match             83.7%;  Score 3527.5;  DB 20;  Length 818;
  Best Local Similarity   83.0%;  Pred. No. 1.8e-304;
  Matches  649;  Conservative   44;  Mismatches   70;  Indels   19;  Gaps    6;

Qy       1 SGPAPRLR----RW------RPNATAGKGVGEV------CAAVVEAATKVEDEGEEDEPV 44
           | | | |     ||      |||  || | | :         | ||    ||: ::|| |
Db      38 ssprpllavpagrwragvrgrpn-vaglgrgrlslhaaaarpvaeavqaeeddddddeev 96

Qy      45 AEDRYALGGACRVLAGMPAPLGATALAGGVNFAVYSGGATAAALCLFTPEDLKADRVTEE 104
           ||:|:|||||||||||||||||||| |||||||| ||:||:| || | ||||||||||
Db      97 aeerfalggacrvlagmpaplgatalrggvnfavyssgasaaslslfapgdlkadrvtee 156

Qy     105 VPLDPLMNRTGNVWHVFIEG-ELHNMLYGYRFDGTFAPHCGHYLDVSNVVVDPYAKAVIS 163
           ||||||:|||||||||||| | ||| || |||||| |||  | | ||||||||||||:|
Db     157 vpldpllnrtgnvwhvfihgdelhgmlcgyrfdgvfapergqyydvsnvvvdpyakavvs 216

Qy     164 RGEYGVPARGNNCWPQMAGMIPLPYSTFDWEGDLPLRYPQKDLVIYEMHLRGFTKHDSSN 223
           |||||||| | :|||||||||||||: |||:|||||  | |||||||||||||||:||
Db     217 rgeygvpapggscwpqmagmiplpynkfdwqgdlplgyhqkdlviyemhlrgftkhnssk 276
```

```
Qy    224 VEHPGTFIGAVSKLDYLKELGVNCIELMPCHEFNELEYSTSSSKMNFWGYSTINFFSPMT 283
          :||||:||||||||:||||||||||||||||||||| :||||||||||||||||||
Db    277 tkhpgtyigavskldhlkelgvncielmpchefneleyfsssskmnfwgystinffspma 336

Qy    284 RYTSGGIKNCGRDAINEFKTFVREAHKRGIEVILDVVFNHTAEGNENGPILSFKGVDNTT 343
          ||:| ||:: |  |||||| |||||||||||||:|||||||||||| ||||||:|:||:|
Db    337 rysssgirdsgcgainefkafvreahkrgievimdvvfnhtaegnekgpilsfrgidnst 396

Qy    344 YYMLAPKGEFYNYSGCGNTFNCNHPVVRQFIVDCLRYWVTEMHVDGFRFDLASIMTRGSS 403
          ||||||||||||||||||||||||||:|||||||||||||||||||||||||:||| |
Db    397 yymlapkgefynysgcgntfncnhpvvrefivdclrywvtemhvdgfrfdlasiltrgcs 456

Qy    404 LWDPVNVYGAPIEGDMITTGTPLVTPPLIDMISNDPILGGVKLIAEAWDAGGLYQVGQFP 463
          |||||||||:|:|||||||||||| ||||||||||||| |||||||||||||| ||||
Db    457 lwdpvnvygspmegdmittgtplvapplidmisndpilgnvkliaeawdagglyqegqfp 516

Qy    464 HWNVWSEWNGKYRDIVRQFIKGTDGFAGGFAECLCGSPHLYQAGGRKPWHSINFVCAHDG 523
          |||||||||||||| |||||||||||| |||||||| |||||||||||| |||||||
Db    517 hwnvwsewngkyrdtvrqfikgtdgfagafaeclcgspqlyqaggrkpwhsigfvcahdg 576

Qy    524 FTLADLVTYNKKYNLPNGENNRDGENHNLSWNCGEEGEFARLSVKRLRKRQMRNFFVCLM 583
          |||||||||| |||| ||||: ||||||||||||||||||| |||:|||||||||||||
Db    577 ftladlvtynskynlsngedfrdgenhnlswncgeegefaslsvrrlrkrqmrnffvclm 636

Qy    584 VSQGVPMFYMGDEYGHTKGGNNNTYCHDSYVNYFRWDKK-EQYSELHRFCCLMTKFRKEC 642
          |||||||||||||||||||||||||| |||||||||| || |:|:||| |||:|||||
Db    637 vsqgvpmfymgdeyghtkggnnntychdhyvnyfrwdkkeeqssdlyrfcrlmtefrkec 696

Qy    643 EGLGLEDFPTAKRLQWHGHQPGKPDWSENSRFVAFSMKDERQGEIYVAFNTSHLPAVVEL 702
          | |||||||::||:|||||||||||||| ||||||:|||| :|||||||||||| || |
Db    697 eslgledfptserlkwhghqpgkpdwseasrfvaftmkdetkgeiyvafntshlpvvvgl 756

Qy    703 PERAGRRWEPVVDTGKPAPYDFLTDDLPDRALTIHQFSHFLYSNLYPMLSYSSVILVLRP 762
          |||:| |||||||||| |||||||| |||||:|::|||||| ||||||||||:||||||
Db    757 persgfrwepvvdtgkeapydfltdglpdravtvyqfshflnsnlypmlsyssiilvlrp 816

Qy    763 DV 764
          ||
Db    817 dv 818
```

RESULT 3
AAY00869
ID AAY00869 standard; Protein; 793 AA.
XX
AC AAY00869;
XX
DT 21-MAY-1999 (first entry)
XX
DE S. tuberosum isoamylase clone 15 protein sequence.
XX
KW Isoamylase clone; potato; branched polysaccharide synthesis; amylopectin;
KW starch modification; debranching activity.
XX
OS Solanum tuberosum.
XX
PN WO9906575-A1.
XX
PD 11-FEB-1999.
XX
PF 30-JUL-1998; 98WO-GB02280.
XX
PR 31-JUL-1997; 97GB-0016185.
XX
PA (PLAN-) PLANT BIOSCIENCE LTD.
XX
PI Bustos Guillen R, Edwards EA, Martin CR, Smith AM;
XX
DR WPI; 1999-153803/13.
DR N-PSDB; AAX27061.
XX
PT New nucleic acid from potato encoding starch debranching enzyme with
PT isoamylase activity - and related vectors, transformed cells and
PT plants, proteins and antibodies, used to generate starch with
PT altered properties, for use in foods, as thickeners etc.
XX
PS Claim 2; Fig 5; 78pp; English.
XX
CC This sequence is a Solanum tuberosum isoamylase of the invention.
CC Fragments of the isoamylase coding sequence are used to identify and
CC clone isoamylases from other plant species, by standard hybridisation
CC or amplification methods. Expression of the isoamylase DNA in host cells
CC is used to produce the isoamylase which is used: (i) to raise antibodies
CC (Ab) (for identification, isolation and localisation of isoamylases); and
CC (ii) for synthesis of branched polysaccharides. The DNA, proteins
CC and Ab (or related peptides) are used to alter the quality and
CC quantity of polysaccharides in a host cells, specifically to alter
CC branching in amylopectin. The modified starches produced are useful in
CC human or animal foods; as biodegradable plastic; as food or paint
CC thickener; in starch-coated films, paper and textiles; in mining
CC explosives; in pharmaceuticals and glues. Manipulation of debranching
CC activity in a plant allows control of starch properties, e.g. increased
CC gel strength; formation of paste rather than gel, changed physical
CC characteristics etc.
XX
SQ Sequence 793 AA;

Query Match 71.9%; Score 3029; DB 20; Length 793;
Best Local Similarity 71.4%; Pred. No. 3.7e-260;
Matches 546; Conservative 78; Mismatches 113; Indels 28; Gaps 7;

```
Qy     10 WRP------NATAGKGVGEVCAAVVEAATKVEDEGEEDEPVAEDRYALGGACR--VLAGM 61
          ||       ||    | | |   |  ||| | |    :| |         || ||:|
Db     37 wrksrssvvnaavdsgrggv---vktaatavvve----kptte-------rcrfevlsgk 82

Qy     62 PAPLGATALAGGVNFAVYSGGATAAALCLFTPEDLKADRVTEEVPLDPLMNRTGNVWHVF 121
          | | |||| |||||||:|  |||| ||| |  ||   ||||:: |||| |:||:|||||
Db     83 plpfgatatdggvnfavfsrnataatlclitlsdlpekrvteqifldplanktgdvwhvf 142

Qy    122 IEGELHNMLYGYRFDGTFAPHCGHYLDVSNVVVDPYAKAVISRGEYGVPARGNNCWPQMA 181
          ::|:  ||||||:||| | |  ||| | | :|:||||||::|||||||    ::||| ||
Db    143 lkgdfenmlygykfdgkfcpeeghyfdssqivldpyakaivsrgeygvlgpeddcwppma 202

Qy    182 GMIPLPYSTFDWEGDLPLRYPQKDLVIYEMHLRGFTKHDSSNVEHPGTFIGAVSKLDYLK 241
          ||:|     |||||||||::||:||||||||:|||| |:||  ::|||::| | |||:||
Db    203 gmvpsasdqfdwegdlplkfpqrdlviyemhvrgftnhessetkypgtylgvvekldhlk 262

Qy    242 ELGVNCIELMPCHEFNELEYSTSSS-----KMNFWGYSTINFFSPMTRYTSGGIKNCGRD 296
          |||||||||||||||||| : :|     | |||||||:|||||| ||:| |: |||
Db    263 elgvncielmpchefneleyysynsvlgdykfnfwgystvnffspmgryssaglsncglg 322

Qy    297 AINEFKTFVREAHKRGIEVILDVVFNHTAEGNENGPILSFKGVDNTTYYMLAPKGEFYNY 356
          ||||||  |:|||||||||:|||||||||||||||||||:|:||: :| ||||||||||
Db    323 ainefkylvkeahkrgievimdvvfnhtaegnengpilsfrgidnsvfytlapkgefyny 382

Qy    357 SGCGNTFNCNHPVVRQFIVDCLRYWVTEMHVDGFRFDLASIMTRGSSLWDPVNVYGAPIE 416
          ||||||||||:|:|||||||||||||||||||||||||:|| || |: ||||| |:
Db    383 sgcgntfncnnpivrqfivdclrywvtemhvdgfrfdlasiltrsssswnavnvygnsid 442

Qy    417 GDMITTGTPLVTPPLIDMISNDPILGGVKLIAEAWDAGGLYQVGQFPHWNVWSEWNGKYR 476
          ||:|||||| :|||||||||||| |||||||||| |||||| |||| :|||||||||
Db    443 gdvittgtpltspplidmisndpilrgvkliaeawdcgglyqvgmfphwgiwsewngkyr 502

Qy    477 DIVRQFIKGTDGFAGGFAECLCGSPHLYQAGGRKPWHSINFVCAHDGFTLADLVTYNKKY 536
          |:|||||||||:| |||||||||:||| ||||||:|||||||||||||||||||| |:
Db    503 dmvrqfikgtdgfsgafaeclcgspnlyqkggrkpwnsinfvcahdgftladlvtynnkh 562

Qy    537 NLPNGENNRDGENHNLSWNCGEEGEFARLSVKRLRKRQMRNFFVCLMVSQGVPMFYMGDE 596
          || |||:|:|||||| |||||||||| : ||:|||||||||:|||||||||| |||||
Db    563 nlangednkdgenhnnswncgeegefasifvkklrkrqmrnfflclmvsqgvpmiymgde 622

Qy    597 YGHTKGGNNNTYCHDSYVNYFRWDKK-EQYSELHRFCCLMTKFRKECEGLGLEDFPTAKR 655
          |||||||||||||:|:|||||||| |  |:  ||| |||||| ||| |||: ||||:|
Db    623 yghtkggnnntychdnyinyfrwdkkdesssdflrfcglmtkfrheceslgldgfptaer 682
```

```
Qy    656 LQWHGHQPGKPDWSENSRFVAFSMKDERQGEIYVAFNTSHLPAVVELPERAGRRWEPVVD 715
          |||||| |  ||||| ||||||:: |: :||:|:||| ||||  : ||:| | ||:| ||
Db    683 lqwhghtprtpdwsetsrfvaftlvdkvkgelyiafnashlpvtitlpdrpgyrwqpfvd 742

Qy    716 TGKPAPYDFLTDDLPDRALTIHQFSHFLYSNLYPMLSYSSVILVL 760
          ||||||:||||||:|:|     |:|||| :| |||||||:||:|
Db    743 tgkpapfdfltddvperetaakqyshfldanqypmlsyssiilll 787


RESULT   4
AAY00872
ID   AAY00872 standard; Protein; 931 AA.
XX
AC   AAY00872;
XX
DT   21-MAY-1999  (first entry)
XX
DE   Original S. tuberosum isoamylase clone 15 protein sequence.
XX
KW   Isoamylase clone; potato; branched polysaccharide synthesis; amylopectin;
KW   starch modification; debranching activity.
XX
OS   Solanum tuberosum.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1..931
FT                   /note= "Xaa= unspecified amino acid"
XX
PN   WO9906575-A1.
XX
PD   11-FEB-1999.
XX
PF   30-JUL-1998;   98WO-GB02280.
XX
PR   31-JUL-1997;   97GB-0016185.
XX
PA   (PLAN-) PLANT BIOSCIENCE LTD.
XX
PI   Bustos Guillen R,  Edwards EA,  Martin CR,  Smith AM;
XX
DR   WPI; 1999-153803/13.
DR   N-PSDB; AAX27066.
XX
PT   New nucleic acid from potato encoding starch debranching enzyme with
PT   isoamylase activity - and related vectors, transformed cells and
PT   plants, proteins and antibodies, used to generate starch with
PT   altered properties, for use in foods, as thickeners etc.
XX
PS   Disclosure; Fig 14; 78pp; English.
XX
CC   This sequence represents the original isolated clone of a
CC   Solanum tuberosum isoamylase of the invention.
CC   Fragments of the isoamylase coding sequence are used to identify and
CC   clone isoamylases from other plant species, by standard hybridisation
CC   or amplification methods. Expression of the isoamylase DNA in host cells
CC   is used to produce the isoamylase which is used: (i) to raise antibodies
CC   (Ab) (for identification, isolation and localisation of isoamylases); and
CC   (ii) for synthesis of branched polysaccharides. The DNA, proteins
CC   and Ab (or related peptides) are used to alter the quality and
CC   quantity of polysaccharides in a host cells, specifically to alter
CC   branching in amylopectin. The modified starches produced are useful in
CC   human or animal foods; as biodegradable plastic; as food or paint
CC   thickener; in starch-coated films, paper and textiles; in mining
CC   explosives; in pharmaceuticals and glues. Manipulation of debranching
CC   activity in a plant allows control of starch properties, e.g. increased
CC   gel strength; formation of paste rather than gel, changed physical
CC   characteristics etc.
XX
SQ   Sequence   931 AA;


  Query Match            69.3%;  Score 2919.5;  DB 20;  Length 931;
  Best Local Similarity  69.6%;  Pred. No. 2.6e-250;

  Matches  534;  Conservative   79;  Mismatches  123;  Indels   31;  Gaps   10;

Qy     10 WRP------NATAGKGVGEVCAAVVEAATKVEDEGEEDEPVAEDRYALGGACRV--LAGM 61
          ||      ||    | | |   |  ||| |  |    :|  |         ||:    |
Db     81 wrksrssvvnaavdsgrggv---vktaatavvve----kptte-------rcrlrfyqgk 126

Qy     62 PAPLGATALAGGVNFAVYSGGATAAALCLFTPEDLKADRVTEEVPLDPLMNRTGNVWHVF 121
          | | ||||  |||||||: | |||| ||| |  ||   ||||:: |||| |:||:|||||
Db    127 plpfgatatdggvnfavfqgnataatlclitlsdlpekrvteqifldplanktgdvwhvf 186

Qy    122 IEGELHNMLYGYRFDGTFAPHCGHYLDVSNVVVDPYAKAVISRGEYGVPARGNNCWPQMA 181
          ::|:  |||||:||| | |  ||| | | :|:||||||::|||||||    ::||| ||
Db    187 lkgdfenmlygykfdgkfcpeeghyfdssqivldpyakaivsrgeygvlgpeddcwppma 246

Qy    182 GMIPLPY-STFDWEGDLPLRYPQKD-LVIYEMHLRGFTKHDSSNVEHPGTFIGAVSKLDY 239
          ||:|  :  :  |||||||    :: | | |||:|||| |:||  ::|||::| | |||:
Db    247 gmvpfcfwislywegdlplevshreilxixemhvrgftihessetkypgtylgvvekldh 306

Qy    240 LKELGVNCIELMPCHEFNELEYSTSSS-----KMNFWGYSTINFFSPMTRYTSGGIKNCG 294
          |||||||||||||||||||||| : :|     | |||||||:|||||| ||:| |: |||
Db    307 lkelgvncielmpchefneleyysynsvlgdykfnfwgystvnffspmgryssaglsncg 366

Qy    295 RDAINEFKTFVREAHKRGIEVILDVVFNHTAEGNENGPILSFKGVDNTTYYMLAPKGEFY 354
            ||||||  |:||||||||||:|||||||||||||||||||:|:||: :| ||||||||
Db    367 lgainefkylvkeahkrgievimdvvfnhtaegnengpilsfrgidnsvfytlapkgefy 426

Qy    355 NYSGCGNTFNCNHPVVRQFIVDCLRYWVTEMHVDGFRFDLASIMTRGSSLWDPVNVYGAP 414
          ||||||||||||:|:|||||  |||||||||| ||||||||:|| || |: |||||
Db    427 nysgcgntfncnnpivrqfiv-mlrywvtemhvhgfrfdlasiltrsssswnavnvygns 485

Qy    415 IEGDMITTGTPLVTPPLIDMISNDPILGGVKLIAEAWDAGGLYQVGQFPHWNVWSEWNGK 474
          |:||:|||||| :|||||||||||| |||||||||| |||||| |||| :|||||||
Db    486 idgdvittgtpltspplidmisndpilrgvkliaeawdcgglyqvgmfphwgiwsewngk 545

Qy    475 YRDIVRQFIKGTDGFAGGFAECLCGSPHLYQAGGRKPWHSINFVCAHDGFTLADLVTYNK 534
          |||:||||||||||:| |||||||||:||| ||||||:||||||||||||||||||||
Db    546 yrdmvrqfikgtdgfsgafaeclcgspnlyqkggrkpwnsinfvcahdgftladlvtynn 605

Qy    535 KYNLPNGENNRDGENHNLSWNCGEEGEFARLSVKRLRKRQMRNFFVCLMVSQGVPMFYMG 594
          |:|| |||:|:|||||| |||||||||| : ||:|||||||||:|||||||||| |||
Db    606 khnlangednkdgenhnnswncgeegefasifvkklrkrqmrnfflclmvsqgvpmiymg 665

Qy    595 DEYGHTKGGNNNTYCHDSYVNYFRWDKK-EQYSELHRFCCLMTKFRKECEGLGLEDFPTA 653
          |||||||||||||||:|:|||||||| |  |:  ||| |||||| ||| |||: ||||
Db    666 deyghtkggnnntychdnyinyfrwdkkdesssdflrfcglmtkfrheceslgldgfpta 725

Qy    654 KRLQWHGHQPGKPDWSENSRFVAFSMKDERQGEIYVAFNTSHLPAVVELPERAGRRWEPV 713
          :|||||| |  ||||| ||||||:: |: :||:|:||| ||||  : ||:| | ||:|
Db    726 erlqwhghtprtpdwsetsrfvaftlvdkvkgelyiafnashlpvtitlpdrpgyrwqpf 785

Qy    714 VDTGKPAPYDFLTDDLPDRALTIHQFSHFLYSNLYPMLSYSSVILVL 760
          |||||||:|||||:|:|     |:|||| :| |||||||:||:|
Db    786 vdtgkpapfdfltddvperetaakqyshfldanqypmlsyssiilll 832


RESULT   5
AAY06918
ID   AAY06918 standard; Protein; 568 AA.
XX
AC   AAY06918;
XX
DT   05-JUL-1999  (first entry)
XX
DE   Partial hexaploid wheat DBE sequence.
XX
KW   Starch biosynthetic pathway; cereal plant; enzyme; SBE; SSS; DBE; GBSS;
KW   starch branching enzyme; starch soluble synthase; debranching enzyme;
KW   endosperm; wheat; barley; granule-bound synthase; glutenin; starch;
KW   grain softness protein I; bacterial isoamylase; glycogen synthase;
KW   wSBE I-D4 gene.
XX
OS   Triticum tauschii.
XX
```

```
FH   Key             Location/Qualifiers
FT   Misc-difference 533
FT                   /label= unknown
FT                   /note= "encoded by TGA"
FT   Misc-difference 541
FT                   /label= unknown
FT                   /note= "encoded by TAA"
FT   Misc-difference 546
FT                   /label= unknown
FT                   /note= "encoded by TAG"
XX
PN   WO9914314-A1.
XX
PD   25-MAR-1999.
XX
PF   11-SEP-1998;   98WO-AU00743.
XX
PR   20-MAR-1998;   98AU-0002509.
PR   12-SEP-1997;   97AU-0009108.
XX
PA   (CSIR ) COMMONWEALTH SCI & IND RES ORG.
PA   (GOOD-) GOODMAN FIELDER LTD.
PA   (LIMA-) GRP LIMAGRAIN PACIFIC PTY LTD.
PA   (AUSU ) UNIV AUSTRALIAN NAT.
XX
PI   Li Z,  Morell M,  Rahman S;
XX
DR   WPI; 1999-229525/19.
DR   N-PSDB; AAX34654.
XX
PT   New isolated cereal plant enzyme genes used for, e.g. expression of
PT   antisense sequences of granule bound synthase
XX
PS   Example 23; Page 100-102; 171pp; English.
XX
CC   The invention relates to a novel enzyme of starch biosynthetic pathway
CC   in a cereal plant, where the enzyme is selected from starch branching
CC   enzyme (SBE) I , SBE II, starch soluble synthase (SSS) I, and debranching
CC   enzyme (DBE), with the proviso that the enzyme is not SSS I of rice, or
CC   SBE I of rice or maize. The methods and products can be used for
CC   targeting expression specifically to the endosperm of the seeds of cereal
CC   plants such as wheat or barley. They can be used for the expression of
CC   e.g. antisense sequences of granule-bound synthase (GBSS), SBE II, low
CC   mol. wt. glutenin, grain softness protein I, bacterial isoamylase,
CC   bacterial glycogen synthase, and wheat high mol. wt. glutenin Bx17. They
CC   can be used for modifying the characteristics of starch produced by a
CC   plant. The present sequence represents the wheat DBE partial sequence.
XX
SQ   Sequence   568 AA;
```

```
  Query Match             69.3%;  Score 2918;  DB 20;  Length 568;
  Best Local Similarity   98.7%;  Pred. No. 1.7e-250;
  Matches  525;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;

Qy     233 AVSKLDYLKELGVNCIELMPCHEFNELEYSTSSSKMNFWGYSTINFFSPMTRYTSGGIKN 292
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 avskldylkelgvncielmpchefneleystssskmnfwgystinffspmtrytsggikn 60

Qy     293 CGRDAINEFKTFVREAHKRGIEVILDVVFNHTAEGNENGPILSFKGVDNTTYYMLAPKGE 352
           ||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||
Db      61 cgrdainefktfvreahkrgievildvvfnhtaegnengpilsfrgvdnttyymlapkge 120

Qy     353 FYNYSGCGNTFNCNHPVVRQFIVDCLRYWVTEMHVDGFRFDLASIMTRGSSLWDPVNVYG 412
           |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db     121 fynysgcgntfncnhpvvrqfivdclrywvtemhvdgfrfdlasimtrgsslwdpvnvyg 180

Qy     413 APIEGDMITTGTPLVTPPLIDMISNDPILGGVKLIAEAWDAGGLYQVGQFPHWNVWSEWN 472
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 apiegdmittgtplvtpplidmisndpilggvkliaeawdagglyqvgqfphwnvwsewn 240

Qy     473 GKYRDIVRQFIKGTDGFAGGFAECLCGSPHLYQAGGRKPWHSINFVCAHDGFTLADLVTY 532
           |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db     241 gkyrdivrqfikgtdgfaggfaeclcgsphlyqaggrkpwhsinfvcahdgftlgdlvty 300

Qy     533 NKKYNLPNGENNRDGENHNLSWNCGEEGEFARLSVKRLRKRQMRNFFVCLMVSQGVPMFY 592
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 nnkynlpngennrdgenhnlswncgeegefarlsvkrlrkrqmrnffvclmvsqgvpmfy 360

Qy     593 MGDEYGHTKGGNNNTYCHDSYVNYFRWDKKEQYSELHRFCCLMTKFRKECEGLGLEDFPT 652
           ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db     361 mgdeyghtkggnnntychdsyvnyfrwdkkeqysdlhrfcclmtkfrkeceglgledfpt 420

Qy     653 AKRLQWHGHQPGKPDWSENSRFVAFSMKDERQGEIYVAFNTSHLPAVVELPERAGRRWEP 712
           |:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 aerlqwhghqpgkpdwsensrfvafsmkderqgeiyvafntshlpavvelperagrrwep 480

Qy     713 VVDTGKPAPYDFLTDDLPDRALTIHQFSHFLYSNLYPMLSYSSVILVLRPDV 764
           |||||||||||||||||||||||||||| ||||||||||||||||||||||
Db     481 vvdtgkpapydfltddlpdraltihqfshflnsnlypmlsyssvilvlrpdv 532
```

```
RESULT   6
AAW36602
ID   AAW36602 standard; Protein; 606 AA.
XX
AC   AAW36602;
XX
DT   08-MAY-1998  (first entry)
XX
DE   S. tuberosum debranching enzyme.
XX
KW   Debranching enzyme; PDE; potato; transgenic plant; starch; industry;
KW   food manufacture.
XX
OS   Solanum tuberosum.
XX
FH   Key             Location/Qualifiers
FT   Protein         1..606
FT                   /note= "partial protein sequence"
XX
PN   DE19618125-A1.
XX
PD   13-NOV-1997.
XX
PF   06-MAY-1996;   96DE-1018125.
XX
PR   06-MAY-1996;   96DE-1018125.
XX
PA   (PLAN-) PLANTEC BIOTECHNOLOGIE GMBH.
XX
DR   WPI; 1997-551167/51.
DR   N-PSDB; AAT96754.
XX
PT   DNA encoding potato debranching enzyme - useful for producing
PT   recombinant enzyme or transgenic plants for use in food production
PT   or in industry
XX
PS   Claim 1; Page 18-19; 20pp; German.
XX
CC   This sequence represents a novel potato debranching enzyme (PDE).
CC   The PDE gene can be used in the production of transgenic plants
CC   especially starch storing plants e.g. Cereal or potato plants. Other
CC   nucleic acid molecules could be designed which specifically hybridise
CC   to the PDE transcript and inhibit its translation. Vectors containing
CC   this gene can be linked in sense orientation with regulatory elements
CC   that ensure its transcription and translation in prokaryotic and
CC   eukaryotic cells. The starch produced from such transgenic plants could
CC   be used in the manufacture of foods or industrial products.
XX
SQ   Sequence   606 AA;
```

```
  Query Match             62.0%;  Score 2612;  DB 18;  Length 606;
  Best Local Similarity   76.6%;  Pred. No. 3e-223;
  Matches  458;  Conservative   62;  Mismatches   70;  Indels    8;  Gaps    3;
```

```
Qy    171 ARG--NNCWPQMAGMIPLPYSTFDWEGDLPLRYPQKDLVIYEMHLRGFTKHDSSNVEHPG 228
Db      3 argpeddcwppmagmvpsasdqfdwegdlllkfpqrdlviyemhvrgftnhessetkypg 62

Qy    229 TFIGAVSKLDYLKELGVNCIELMPCHEFNELEYSTSSS-----KMNFWGYSTINFFSPMT 283
Db     63 tylgvvekldhlkelgvncielmpchefneleyysynsvlgdykfnfwgystvnffspmg 122

Qy    284 RYTSGGIKNCGRDAINEFKTFVREAHKRGIEVILDVVFNHTAEGNENGPILSFKGVDNTT 343
Db    123 ryssaglsncglgainefkylvkeahkrgievimdvvfnhtaegnengpilsfrgidnsv 182

Qy    344 YYMLAPKGEFYNYSGCGNTFNCNHPVVRQFIVDCLRYWVTEMHVDGFRFDLASIMTRGSS 403
Db    183 fytlapkgefynysgcgntfncnnpivrqfivdclrywvtemhvdgfrfdlasiltrsss 242

Qy    404 LWDPVNVYGAPIEGDMITTGTPLVTPPLIDMISNDPILGGVKLIAEAWDAGGLYQVGQFP 463
Db    243 swnavnvygnsidgdmittgtpltspplidmisndpilsgvkliaeawdcgglyqvgmfp 302

Qy    464 HWNVWSEWNGKYRDIVRQFIKGTDGFAGGFAECLCGSPHLYQAGGRKPWHSINFVCAHDG 523
Db    303 hwgiwsewngkyrdmvrqfikgtdgfsgafaeclcgspnlyqkggrkpwnsinfvcahdg 362

Qy    524 FTLADLVTYNKKYNLPNGENNRDGENHNLSWNCGEEGEFARLSVKRLRKRQMRNFFVCLM 583
Db    363 ftladlvtynnkhnlangednkdgenhnnswncgeegefasifvkklrkrqmrnfflclm 422

Qy    584 VSQGVPMFYMGDEYGHTKGGNNNTYCHDSYVNYFRWDKK-EQYSELHRFCCLMTKFRKEC 642
Db    423 vsqgvpmiymgdeyghtkggnnntychdnyinyfrwdkkdesssdflrfcglmtkfrhec 482

Qy    643 EGLGLEDFPTAKRLQWHGHQPGKPDWSENSRFVAFSMKDERQGEIYVAFNTSHLPAVVEL 702
Db    483 eslgldgfptaerlqwhghtprtpdwsetsrfvaftlvdkvkgelyiafnashlpvtitl 542

Qy    703 PERAGRRWEPVVDTGKPAPYDFLTDDLPDRALTIHQFSHFLYSNLYPMLSYSSVILVL 760
Db    543 pekpgyrwqpfvdtgkpapfdfltddvperetaakqyshfldanqypmlsyssiilll 600
```

```
RESULT   7
AAY04382
ID   AAY04382 standard; Protein; 429 AA.
XX
AC   AAY04382;
XX
DT   06-JUL-1999  (first entry)
XX
DE   Potato isoamylase type DBE genomic clone protein SEQ ID NO:15.
XX
KW   Potato; isoamylase debranching enzyme; DBE; starch.
XX
OS   Solanum tuberosum.
XX
PN   WO9912950-A2.
XX
PD   18-MAR-1999.
XX
PF   04-SEP-1998;   98WO-GB02665.
XX
PR   06-SEP-1997;   97GB-0018863.
XX
PA   (NATT ) NAT STARCH & CHEM INVESTMENT HOLDING COR.
XX
PI   Jobling SA,  Schwall GP,  Westcott RJ;
XX
DR   WPI; 1999-229220/19.
DR   N-PSDB; AAX33162.
XX
PT   New isolated potato isoamylase-type debranching enzyme gene
XX
PS   Claim 6; Fig 8; 72pp; English.
XX
CC   The present invention describes a novel nucleic acid sequence which is
CC   obtainable from potato plants and carries at least a portion of an
CC   isoamylase-type debranching enzyme (DBE) gene. Constructs containing the
CC   nucleic acid sequence can be used to alter the starch properties of
CC   plants such as potato, sweet potato, maize, wheat, barley, oat, cassava,
CC   pea or rice. The starch can have increased branching and/or shorter
CC   chain length, reduced peak viscosity, higher setback viscosity or
CC   increased viscosity onset temperature. By using an antisense sequence
CC   with greater homology to the native gene, greater inhibition can be
CC   achieved. The present sequence represents a protein from an isoamylase
CC   type DBE genomic clone.
XX
SQ   Sequence   429 AA;
```

```
  Query Match             44.7%;  Score 1881;  DB 20;  Length 429;
  Best Local Similarity   77.4%;  Pred. No. 1.8e-158;
  Matches  328;  Conservative   44;  Mismatches   50;  Indels    2;  Gaps    2;

Qy    338 GVDNTTYYMLAPKGEFYNYSGCGNTFNCNHPVVRQFIVDCLRYWVTEMHVDGFRFDLASI 397
Db      1 gidnsvfytlapkgefynysgcgntfncnnpivrqfivdclrywvtemhvdgfrfdlasi 60

Qy    398 MTRGSSLWDPVNVYGAPIEGDMITTGTPLVTPPLIDMISNDPILGGVKLIAEAWDAGGLY 457
Db     61 ltrsssswnavnvygnsidgdvittgtpltspplidmisndpilrgvkliaeawdcggly 120

Qy    458 QVGQFPHWNVWSEWNGKYRDIVRQFIKGTDGFAGGFAECLCGSPHLYQAGGRKPWHSINF 517
Db    121 qvgmfphwgiwsewn-evrdmvrqfidgtdgfsgafaeclcgspnlyqkggrkpwnsinf 179

Qy    518 VCAHDGFTLADLVTYNKKYNLPNGENNRDGENHNLSWNCGEEGEFARLSVKRLRKRQMRN 577
Db    180 vcahdgftladlvtynnkhnlangednkdgenhnnswncgeegefasifmkklrkrqmrn 239

Qy    578 FFVCLMVSQGVPMFYMGDEYGHTKGGNNNTYCHDSYVNYFRWDKK-EQYSELHRFCCLMT 636
Db    240 fflclmvsqgvpmiymgdeyghtkggnnntychdnyinyfrwdkkdesssdflrfcglmt 299

Qy    637 KFRKECEGLGLEDFPTAKRLQWHGHQPGKPDWSENSRFVAFSMKDERQGEIYVAFNTSHL 696
Db    300 kfrheceslgldgfptaerlqwhghtprtpdwsetsrfvaftlvdkvkgelyiafnashl 359

Qy    697 PAVVELPERAGRRWEPVVDTGKPAPYDFLTDDLPDRALTIHQFSHFLYSNLYPMLSYSSV 756
Db    360 pvtitlpdrpgyrwqpfvdtgkpapfdfltddvperetaakqyshfldanqypmlsyssi 419

Qy    757 ILVL 760
Db    420 illl 423
```

```
RESULT   8
AAY00870
ID   AAY00870 standard; Protein; 766 AA.
XX
AC   AAY00870;
XX
DT   21-MAY-1999  (first entry)
XX
DE   S. tuberosum isoamylase clone 9 protein sequence.
XX
KW   Isoamylase clone; potato; branched polysaccharide synthesis; amylopectin;
KW   starch modification; debranching activity.
XX
OS   Solanum tuberosum.
XX
PN   WO9906575-A1.
XX
PD   11-FEB-1999.
XX
```

```
PF   30-JUL-1998;   98WO-GB02280.
XX
PR   31-JUL-1997;   97GB-0016185.
XX
PA   (PLAN-) PLANT BIOSCIENCE LTD.
XX
PI   Bustos Guillen R,  Edwards EA,  Martin CR,  Smith AM;
XX
DR   WPI; 1999-153803/13.
DR   N-PSDB; AAX27062.
XX
PT   New nucleic acid from potato encoding starch debranching enzyme with
PT   isoamylase activity - and related vectors, transformed cells and
PT   plants, proteins and antibodies, used to generate starch with
PT   altered properties, for use in foods, as thickeners etc.
XX
PS   Claim 2; Fig 6; 78pp; English.
XX
CC   This sequence is a Solanum tuberosum isoamylase of the invention.
CC   Fragments of the isoamylase coding sequence are used to identify and
CC   clone isoamylases from other plant species, by standard hybridisation
CC   or amplification methods. Expression of the isoamylase DNA in host cells
CC   is used to produce the isoamylase which is used: (i) to raise antibodies
CC   (Ab) (for identification, isolation and localisation of isoamylases); and
CC   (ii) for synthesis of branched polysaccharides. The DNA, proteins
CC   and Ab (or related peptides) are used to alter the quality and
CC   quantity of polysaccharides in a host cells, specifically to alter
CC   branching in amylopectin. The modified starches produced are useful in
CC   human or animal foods; as biodegradable plastic; as food or paint
CC   thickener; in starch-coated films, paper and textiles; in mining
CC   explosives; in pharmaceuticals and glues. Manipulation of debranching
CC   activity in a plant allows control of starch properties, e.g. increased
CC   gel strength; formation of paste rather than gel, changed physical
CC   characteristics etc.
XX
SQ   Sequence   766 AA;

  Query Match             37.4%;  Score 1577;  DB 20;  Length 766;
  Best Local Similarity   44.0%;  Pred. No. 4.9e-131;
  Matches  329;  Conservative  111;  Mismatches  235;  Indels   72;  Gaps   19;

Qy      30 AATKVEDEGEEDEPVAEDRYALGGACRVLAGMPAPLGATALAGGVNFAVYSGGATAAALC 89
           || |:::|     |   | :    : :|  |:  ||| :    |:|||::|  |:|  ||
Db      70 aaaklqeea----pqmldtfp---sfkvspglahplgvsetesginfaifsqhasavtlc 122

Qy      90 LFTPEDLKADRVTEEVPLDPLMNRTGNVWHVFIEGEL--HNMLYGYRFDGTFAPHCGHYL 147
           :  |: :    :  |: ||| ||||::||: |: ||    :|||| ||    | ||
Db     123 iilpksvhdgmi--elaldpqknrtgdiwhicik-elpqggvlygyridgprnwheghrf 179

Qy     148 DVSNVVVDPYAKAVISRGEYGVPARGNNCWPQMAGMIPLPYSTFDWEGDLPL-RYPQKDL 206
           | | ::|||||| :  |  :|     :|   :  |         ||| :  |   |:|||
Db     180 ddsiilvdpyakliegrrvfg---desnkmcrffgtydfnslpfdwgenyklpnipekdl 236

Qy     207 VIYEMHLRGFTKHDSSNVEHP--GTFIGAVSKLDYLKELGVNCIELMPCHEFNELEY--- 261
           |||||::| ||  ::|:::    |:::| : |: :| ||||| :||:|  ||:|||
Db     237 viyemnvraftadetssldqdqrgsylgliekiphllelgvnavellpvfefdelelqrr 296

Qy     262 -STSSSKMNFWGYSTINFFSPMTRYTSGGIKNCGRDAIN---EFKTFVREAHKRGIEVIL 317
            :     :| ||||||||||:||:|| |     ||   :    |||  |:  |  ||||||
Db     297 pnprdhmintwgystinffapmsryas-----cgggpvraswefkemvkalhgagievil 351

Qy     318 DVVFNHTAEGNENGP-ILSFKGVDNTTYYM--LAPKGEFYNYSGCGNTFNCNHPVVRQFI 374
           |||:||| | ::  |   ||:|:||  ||| |    :  |::|||||||||| | : |
Db     352 dvvynhtneaddenpyttsfrgidnkvyymvdlnnnaqllnfagcgntfncnhptvmeli 411

Qy     375 VDCLRYWVTEMHVDGFRFDLASIMTRGSSLWDPVNVYGAPIEGDMITTGTPLVTPPLIDM 434
           :: ||:|||| ||||||||||:: ||                 | |||:  |||:
Db     412 leslrhwvteyhvdgfrfdlasvlcrg-----------------tdgtpinapplvka 452

Qy     435 ISNDPILGGVKLIAEAWDAGGLYQVGQFPHWNVWSEWNGKYRDIVRQFIKGTDGFAGGFA 494
           || | :|   |:||| || |||| ||:||:|: |:|||||||| :|:|||  |  | ||
Db     453 iskdsvlsrckiiaepwdcgglylvgkfpnwdrwaewngkyrddirrfikgdagmkgnfa 512

Qy     495 ECLCGSPHLYQAGGRKPWHSINFVCAHDGFTLADLVTYNKKYNLPNGENNRDGENHNLSW 554
             : ||  ||:   |||:||:||| |||||| |||:|| |:|  |||   || | | ||
Db     513 triagsadlyrvnkrkpyhsvnfviahdgftlydlvsynnkhndangeggndgcndnfsw 572

Qy     555 NCGEEGEFARLSVKRLRKRQMRNFFVCLMVSQGVPMFYMGDEYGHTKGGNNNTYCHDSYV 614
           ||| ||| :  ::  || |||:|| : |||||| ||  |||||||: ||||:| ||: :
Db     573 ncgiegetsdaninalrsrqmknfhlalmvsqgtpmmlmgdeyghtrygnnnsyghdtai 632

Qy     615 NYFRWDKKEQYSELH-RFCCLMTKFRKECEGLGLEDFPTAKRLQWHGHQPGKPDW-SENS 672
           | |:| : |     | ||   | |||     |  |:|     : |      : :| :| |
Db     633 nnfqwgqlearkndhfrffskmikfrlshnvlrkenfieknditwl-----ednwynees 687

Qy     673 RFVAFSMKDERQGEIYVAFNTSHLPAVVELPE-RAGRRWEPVVDTGKPAPYDFLTDDLPD 731
           ||:|| : |   |:||:||| |      :|     | |  ||||   :| ||:|: :
Db     688 rflafmlhdgnggdiylafnahhfsiktaipspprnrswyrvvdtnlkspddfvtegvsg 747

Qy     732 RALTIHQFSHFLYSNLYPMLSYSSVIL 758
                     |  | :  ||:::|
Db     748 ------------isktydvapysaill 762


RESULT   9
AAY34991
ID   AAY34991 standard; Protein; 666 AA.
XX
AC   AAY34991;
XX
DT   13-SEP-1999  (first entry)
XX
DE   C. pneumoniae protein involved in intermediate metabolism.
XX
KW   Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;
KW   sinusitis; purulent otitis media; erythema nodosum; pharyngitis;
KW   vaccine; neutralising epitope.
XX
OS   Chlamydia pneumoniae.
XX
PN   WO9927105-A2.
XX
PD   03-JUN-1999.
XX
PF   20-NOV-1998;   98WO-IB01890.
XX
PR   04-NOV-1998;   98US-0107078.
PR   21-NOV-1997;   97FR-0014673.
XX
PA   (GEST ) GENSET.
XX
PI   Griffais R;
XX
DR   WPI; 1999-357842/30.
XX
PT   Genome sequence of Chlamydia pneumoniae
XX
PS   Page 898-899; Disclosure; 1912pp; English.
XX
CC   AAY34584-Y35879 represent the proteins encoded by all the open reading
CC   frames in the complete genome (see AAX91990) of Chlamydia pneumoniae.
CC   C. pneumoniae causes respiratory disease such as pneumonia and
CC   bronchitis and is thought to be a contributing factor in heart
CC   disease, sarcoidosis, sinusitis, purulent otitis media, erythema
CC   nodosum or pharyngitis. The polypeptides encoded by the open reading
CC   frames of the C. pneumoniae genome (see AAY34584-Y35879) can be used in
CC   immunogenic compositions as vaccines. Vectors containing C. pneumoniae
CC   nucleotides sequences can also be used as immunogenic compositions,
CC   especially where the vector directs the expression of a neutralising
CC   epitope of C. pneumoniae.
XX
SQ   Sequence   666 AA;

  Query Match             31.0%;  Score 1305.5;  DB 20;  Length 666;
```

```
  Best Local Similarity   42.6%;  Pred. No. 5.5e-107;
  Matches  279;  Conservative   88;  Mismatches  231;  Indels   57;  Gaps   16;

Qy      61 MPAPLGATALA-GGVNFAVYSGGATAAALCLFTPEDLKADRVTE--EVPLDPLMNRTGNV 117
           :| ||||: ::     ||:|:  ||   | |        |  :|  |||| |  :||| :
Db      12 vplplgaskispnryrfalyasqatevilal-------tdensevievplypdthrtgai 64

Qy     118 WHVFIEGELHNMLYGYRFDGTFAPHCGHYLDVSNVVVDPYAKAVISRGEYGVPARGNN-- 175
           ||: |||       | :|  |    | |        : ||||| : |   :|   :  :
Db      65 whieiegisdqssyafrvhgp-kkh-gmqysfkeyladpyaknihspqsfgsrkkqgdya 122

Qy     176 -CWPQMAGMIPLPYSTFDWEGDLPLRYPQKDLVIYEMHLRGFTKHDSSNVEHPGTFIGAV 234
            |:        |    | |:|| ||  |:::::|||||:| ||:  || |  ||||:| :
Db     123 fcy--------lkeepfpwdgdqplhlpkeemiiyemhvrsftqssssrvhapgtflgii 174

Qy     235 SKLDYLKELGVNCIELMPCHEFNELEYSTSSSKM----NFWGYSTINFFSPMTRYTSGGI 290
            |:|:| :||:| :||:|  ||:|  :   :||     |:|||: :|||||  ||
Db     175 ekidhlhklginavellpifefdetahpfrnskfpylcnywgyaplnffspcrryayas- 233

Qy     291 KNCGRDAINEFKTFVREAHKRGIEVILDVVFNHTAEGNENGPILSFKGVDNTTYYMLAPK 350
             |      |||| |:  |: |||||||||||||      |   |   :|  :||:|  :
Db     234 dpcaps--refktlvktlhqegievildvvfnht---glqgttcslpwidtpsyyildaq 288

Qy     351 GEFYNYSGCGNTFNCNHPVVRQFIVDCLRYWVTEMHVDGFRFDLASIMTRGSSLWDPVNV 410
           | | |||||||| | |     |:|:| ||||| ||||||||||||||: :|| |
Db     289 ghftnysgcgntlntnrapttqwildilrywveemhvdgfrfdlasvfsrgps------- 341

Qy     411 YGAPIEGDMITTGTPLVTPPLIDMISNDPILGGVKLIAEAWDAGGLYQVGQFPHWNV-WS 469
                 |:||   |::: || ||:|   |:||| |||||||||| ||  :  ||
Db     342 -----------gsplqfapvleaisfdpllastkiiaepwdagglyqvgyfptlsprws 389

Qy     470 EWNGKYRDIVRQFIKGTDGFAGGFAECLCGSPHLYQAGGRKPWHSINFVCAHDGFTLADL 529
           |||| ||| |: |: |     | ||  : ||  :|  |   | :|||:|  ||||||
Db     390 ewngpyrdnvkaflngdqnligtfasrisgsqdiyphg--sptnsinyvschdgftlcdt 447

Qy     530 VTYNKKYNLPNGENNRDGENHNLSWNCGEEGEFARLSVKRLRKRQMRNFFVCLMVSQGVP 589
           |||| |:|  |||:|||| : | |:| | ||:     :  :|:||:||||: ||||||:|
Db     448 vtynhkhneangednrdgtdanysynfgtegktedpgilevrerqlrnffltlmvsqgip 507

Qy     590 MFYMGDEYGHTKGGNNNTYCHDSYVNYFRWDKKEQYSELHRFCCLMTKFRKECEGLGLED 649
           |   |||| ||  |||| :  ||  ||| ||:       |  | | :  |||: : |
Db     508 miqsgdeyahtaegnnnrwaldsnanyflwdqltakptlmhflcdliafrkkyktlfnrg 567

Qy     650 FPTAKRLQWHGHQPGKPDWSENSRFVAFSMKDERQGEIYVAFNTSHLPAVVELPE 704
           | : | : |          |   : |:|| :|  :   :||||:      :  ||:
Db     568 flsnkeiswvdamgnpmtwrpgn-flafkikspk-ahvyvafhvgaqdqlatlpk 620


RESULT  10
AAY37184
ID   AAY37184 standard; Protein; 618 AA.
XX
AC   AAY37184;
XX
DT   07-OCT-1999  (first entry)
XX
DE   Protein involved in intermediate metabolism of sugars and/or cofactors.
XX
KW   Vaccine; eye disease; conventional trachoma; nonendemic trachoma;
KW   paratrachoma; inclusion conjunctivitis; genital disease; perihepatitis;
KW   nongonococcal uretritis; epidymitis; cervicitis; salpingitis;
KW   bartholinitis; pneumopathy; venereal lymphogranulomatosis.
XX
OS   Chlamydia trachomatis.
XX
PN   WO9928475-A2.
XX
PD   10-JUN-1999.
XX
PF   27-NOV-1998;   98WO-IB01939.
XX
PR   04-NOV-1998;   98US-0107077.
PR   28-NOV-1997;   97FR-0015041.
PR   17-DEC-1997;   97FR-0016034.
XX
PA   (GEST ) GENSET.
XX
PI   Griffais R;
XX
DR   WPI; 1999-371125/31.
XX
PT   Genome sequence of Chlamydia trachomatis
XX
PS   Disclosure; Page 958-959; 1755pp; English.
XX
CC   AAY36754-Y37949 are encoded by open reading frames (ORFs) of the genome
CC   of Chlamydia trachomatis (see AAZ01425). The polypeptides can be used as
CC   vaccines against Chlamydia trachomatis. Antisense and ribozyme sequences
CC   can also be used to control growth of the microorganism. Chlamydia
CC   trachomatis is responsible for a large number of diseases, e.g. eye
CC   diseases such as conventional trachoma, nonendemic trachoma,
CC   paratrachoma, and inclusion conjunctivitis; genital diseases such as
CC   nongonococcal uretritis, epidymitis, cervicitis, salpingitis,
CC   perihepatitis, bartholinitis; pneumopathy in breast feeding infants;
CC   and venereal lymphogranulomatosis. The polypeptides of the invention
CC   may be of use in treating these diseases.
XX
SQ   Sequence   618 AA;


  Query Match             28.6%;  Score 1206.5;  DB 20;  Length 618;
  Best Local Similarity   40.2%;  Pred. No. 3.1e-98;
  Matches  262;  Conservative   92;  Mismatches  208;  Indels   89;  Gaps   17;

Qy     142 HCGHYL----------DVS-------NVVVDPYAKAVISRGEYGVPARGNNCWPQMAG 182
           :| |:|           |:|       : : |||:| : |   :| | :  :
Db      18 ncrhflvngrmlinlrgtdlssqkfatdsyiadpyskniyspqlfgspkqekd-----ya 72

Qy     183 MIPLPYSTFDWEGDLPLRYPQKDLVIYEMHLRGFTKHDSSNVEHPGTFIGAVSKLDYLKE 242
             | |  |||||| ||  |:::  |||||:| ||:  || | |||||:| : |:|:||:
Db      73 fsylkyedfdwegdtplhlpkenyfiyemhvrsftrdpssqvshpgtflgiiekidhlkq 132

Qy     243 LGVNCIELMPCHEFNELEYSTSSSKM----NFWGYSTINFFSPMTRYTSGGIKNCGRDAI 298
           |||: :||:|  ||:|  :   :       |:||||::||| |  ||| |    |
Db     133 lgvhavellpifefdetvhpfknqdfphlcnywgyssvnffcpsrrytyga-dpca--pa 189

Qy     299 NEFKTFVREAHKRGIEVILDVVFNHTAEGNENGPILSFKGVDNTTYYMLAPKGEFYNYSG 358
            |||| |:  |: |||||||||||||     : |:     :|  :|||:   |:  |:||
Db     190 refktlvkalhragievildvvfnhtgfegtscpl---pwidlksyymvndhgdlmnfsg 246

Qy     359 CGNTFNCNHPVVRQFIVDCLRYWVTEMHVDGFRFDLASIMTRGSSLWDPVNVYGAPIEGD 418
           |||| | | |   ::|:| ||||| |||| ||||||||: :|     ||
Db     247 cgntvntntpttlkwildalrywvqemhvfgfrfdlasvfsr-----dp---------- 290

Qy     419 MITTGTPLVTPPLIDMISNDPILGGVKLIAEAWDAGGLYQVGQFPHWNV-WSEWNGKYRD 477
             | ||   |::  ||:| ||   ||||| ||||||||:| ||  :  ||||| |||
Db     291 ---qgvplpltpilqaissdsilsetkliaepwdagglyqlghfpsistrwsewngcyrd 347

Qy     478 IVRQFIKGTDGFAGGFAECLCGSPHLYQAGGRKPWHSINFVCAHDGFTLADLVTYNKKYN 537
            |: |: |       ||  : ||  :|  |  || :|||::|:||||||| | | || |:|
Db     348 hvkaflngdahqvssfasrisgshdiypng--kptnsinyicshdgftlydtvayndkhn 405

Qy     538 LPNGENNRDGENHNLSWNCGEEGEFARLSVKRLRKRQMRNFFVCLMVSQGVPMFYMGDEY 597
             ||| |||| : | |:| | |||     ::  ||:|||:|||: | :|||:||   ||||
Db     406 eengeynrdgtsanysynfgcegettdpticalrerqmknfflalflsqgipmiqsgdey 465

Qy     598 GHTKGGNNNTYCHDSYVNYFRWDKKEQYSELHRFCCLMTKFRKECEGLGLEDFPTAKRLQ 657
           |||  |||| :| |: :||| ||:  :  ||  | | :   ||    |    | :   :
Db     466 ghtaygnnnhwcldtkinyflwdrlaerkelfsflcqvialrkaytelfntsflsedtit 525

Qy     658 WHGHQPGKPDWSENSRFVAFSMKDERQGEIYVAFNTSHLPAVVELPE--RAGRRWEPVVD 715
           |   :    :|  :  ::|| :|       ::| | : :    : ||:  :    :| :||
Db     526 wlntkgsprewgad-hylafxlk-hlnyslfvxfysgnerieislpkprkehlayekivd 583

Qy     716 --TGKPAPYDFLTDDLPDRALTIHQFSHFLYSNLYPMLS---YSSVILVLR 761
             ||                          |    | | ||   |||:: : |
```

```
Db    584 sttg----------------------ffsqilspklslepysslvaisr 610


RESULT  11
AAG92051
ID   AAG92051 standard; Protein; 836 AA.
XX
AC   AAG92051;
XX
DT   26-SEP-2001  (first entry)
XX
DE   C glutamicum protein fragment SEQ ID NO: 5805.
XX
KW   Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
KW   organic acid synthesis.
XX
OS   Corynebacterium glutamicum.
XX
PN   EP1108790-A2.
XX
PD   20-JUN-2001.
XX
PF   18-DEC-2000; 2000EP-0127688.
XX
PR   16-DEC-1999;   99JP-0377484.
PR   07-APR-2000; 2000JP-0159162.
PR   03-AUG-2000; 2000JP-0280988.
XX
PA   (KYOW ) KYOWA HAKKO KOGYO KK.
XX
PI   Nakagawa S,  Mizoguchi H,  Ando S,  Hayashi M,  Ochiai K,  Yokoi H;
PI   Tateishi N,  Senoh A,  Ikeda M,  Ozaki A;
XX
DR   WPI; 2001-376931/40.
DR   N-PSDB; AAH67270.
XX
PT   Novel polynucleotides derived from Coryneform bacteria, for identifying
PT   mutation point of a gene, measuring expression of a gene, analysing
PT   expression profile or pattern of a gene and identifying homologous gene
PT   -
XX
PS   Claim 17; SEQ ID NO: 5805; 246pp + Sequence Listing; English.
XX
CC   The present invention provides a number of nucleotide and protein
CC   sequences from the Coryneform bacterium Corynebacterium glutamicum. These
CC   are useful for identifying the mutation point of a gene derived from a
CC   mutant of coryneform bacterium, measuring expression amount and
CC   analysing the expression profile or expression pattern of a gene derived
CC   from Coryneform bacterium, and identifying a homologue of a gene derived
CC   from coryneform bacterium. Coryneform bacteria are useful for producing
CC   amino acids, nucleic acids, vitamins, saccharides and organic acids,
CC   particularly L-lysine. The present sequence is a protein described
CC   in the exemplification of the invention.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from the
CC   European Patent Office.
XX
SQ   Sequence   836 AA;


  Query Match           28.6%;  Score 1204;  DB 22;  Length 836;
  Best Local Similarity  38.5%;  Pred. No. 8.3e-98;
  Matches  283;  Conservative   98;  Mismatches  240;  Indels  114;  Gaps   25;

Qy     56 RVLAGMPAPLGATALAGGVNFAVYSGGATAAALCLFTPEDLKADRVTEEVPLDPLMNRTG 115
Db     10 qvwpghayplgstydgagtnfalfsdvaervelcl-----ldadnnetriple---erda 61

Qy    116 NVWHVFIEGELHNMLYGYRFDGTFAPHCGHYLDVSNVVVDPYAKAVISRGEY-GVPA--- 171
Db     62 hiwhcylpgvqpgqrygfrvhgpwnpdegkrcdankllvdpyaraf--dgdfdghpslfs 119

Qy    172 ---------RGNNCWPQ----MAGMIPLPYSTFDWEGDLPLRYPQKDLVIYEMHLRGFT- 217
                    | |       |  ::  |: |||  |   | |  : |||| |::| |
Db    120 yditnpndpngrntedsidhtmksvvvnpf--fdwgndrapripynetviyeahvkgmtm 177

Qy    218 KHDSSNVEHPGTFIGAV--SKLDYLKELGVNCIELMPCHEFNELEYSTSSSKMNFWGYST 275
          |     |  ||: |    : : || :||| ||||| |:| : :         |:|||::
Db    178 thpdvpeelrgtyaglahptiiqylsdlgvtaielmpvhqflqddrlrdlgmrnywgyns 237

Qy    276 INFFSPMTRYTSGGIKNCGRDAINEFKTFVREAHKRGIEVILDVVFNHTAEGNENGPILS 335
            ||:|   | :   || |  |: |||  ||  |: |:|||||||:|||||||  || ::
Db    238 fgffapyndyaan--knpg-gavaefkglvrsyheaglevildvvynhtaegnhmgptia 294

Qy    336 FKGVDNTTYYML--APKGEFYNYSGCGNTFNCNHPVVRQFIVDCLRYWVTEMHVDGFRFD 393
          |:|:||  || |     :  : :|:| ||: |   |   | |:| ||||||||||||||||
Db    295 frgidneayyrlvegdrrhymdytgtgnslnvrdphslqlimdslrywvtemhvdgfrfd 354

Qy    394 LASIMTRGSSLWDPVNVYGAPIEGDMITTGTPLVTPPLIDMISNDPILGGVKLIAEAWDA 453
          ||| : |    :| |          | : |          |::  ||::  |||||| ||
Db    355 lastlare---fddv---------drlat--------ffdlvqqdpvvsqvkliaepwdv 394

Qy    454 G-GLYQVGQFPHWNVWSEWNGKYRDIVRQFIKGTDGFAGGFAECLCGSPHLYQAGGRKPW 512
          | | |||| ||   :|:|||||||| || | :|     | ||  | ||  ||   ||:|
Db    395 geggyqvgnfp--plwtewngkyrdtvrdfwrgepatlgefasrltgssdlyanngrrpt 452

Qy    513 HSINFVCAHDGFTLADLVTYNKKYNLPNGENNRDGENHNLSWNCGEEGEFARLSVKRLRK 572
          |||||  |||||||  |||:||:|:|: |||: ||||:|| |||||  ||       : :||
Db    453 asinfvtahdgftlndlvsynekhnmangedgrdgeshnrswncgvegptddpeimqlra 512

Qy    573 RQMRNFFVCLMVSQGVPMFYMGDEYGHTKGGNNNTYCHDSYVNYFRWDKKEQYSELHRFC 632
          :| |||   |::||| ||   |||   |: |||| || |: : :  ||: |: ::|  |
Db    513 qqrrnflttlllsqgtpmlshgdemartqngnnnvycqdnelawvnwdqaeenadlvsft 572

Qy    633 CLMTKFRKECE----------GLGLE--------------------DFPTAKRLQ--W 658
            : : |               || :                    ||   | ||  :
Db    573 rrllriranhpvfrrrqflaggplgadvrdrdiawlvpngtlmtqddwdfafgkslqvff 632

Qy    659 HGHQPGKPDW-----SENSRFVAFSMKDERQGEIYVAFNTSHLPAVVELPERAGRRWEPV 713
          :|    :||:      ::|  : |:   |      : ||   ||     ||  | :|: :
Db    633 ngdaieepdyrgqkihddsfilmfnahfep-----idfn---lp-----pehfgmkwkll 679

Qy    714 VDTGKPA--PYDFLT 726
          ||| :    | : ||
Db    680 vdtteavghpledlt 694


RESULT  12
AAW95636
ID   AAW95636 standard; Protein; 726 AA.
XX
AC   AAW95636;
XX
DT   08-JUN-1999  (first entry)
XX
DE   Rhodothermus marinus DSM 4252 isoamylase.
XX
KW   Isoamylase; starch conversion.
XX
OS   Rhodothermus marinus.
XX
PN   WO9901545-A1.
XX
PD   14-JAN-1999.
XX
PF   02-JUL-1998;   98WO-DK00304.
XX
PR   13-AUG-1997;   97US-0055867.
PR   02-JUL-1997;   97DK-0000787.
XX
PA   (NOVO ) NOVO-NORDISK AS.
XX
PI   Bisgard-Frantzen H,  Svendsen A,  Tsutsumi N;
XX
DR   WPI; 1999-106047/09.
DR   N-PSDB; AAX07446.
```

XX
PT Starch conversion processes - using iso:amylase(s), particularly
PT obtained using novel DNA cloned from Rhodothermus or Sulfolobus
PT strains
XX
PS Claim 23; Pages 67-69; 76pp; English.
XX
CC The sequence is that of an isoamylase which can be used as part of a
CC method of starch conversion for the production of high fructose
CC syrup or fat replacers. The use of an isoamylase in starch conversion
CC results in a reduced formation of the undesired trisaccharide panose
CC to increase the saccharification yield.
XX
SQ Sequence 726 AA;

Query Match 28.2%; Score 1187; DB 20; Length 726;
Best Local Similarity 38.9%; Pred. No. 2.2e-96;
Matches 294; Conservative 86; Mismatches 242; Indels 134; Gaps 30;

Qy 57 VLAGMPAPLGATALAGGVNFAVYSGGATAAALCLFT-PEDLKADRVTEEVPLDPLMNRTG 115
Db 14 vwpgrpyplgatwdglgvnfalysqhaeavelvlfdhpddpapsrtie------vtertg 67

Qy 116 NVWHVFIEGELHNMLYGYRFDGTFAPHCGHYLDVSNVVVDPYAKAVISRGEYGVPARGNN 175
Db 68 piwhvylpglrpgqlygyrvygpyrpeeghrfnpnkvlldpyakai------grplr--- 118

Qy 176 CWPQM---------AGMIPL------PYS--------TFDWEGDLPLRYPQKDLVIYEMH 212
Db 119 -whdslfgykigdpagdlsfseedsapyaplgavvegcfewgddrppripwedtiiyeth 177

Qy 213 LRGFTKHDSSNVEHP-------GTFIGAVSK--LDYLKELGVNCIELMPCHEFNELEYST 263
Db 178 vkgitk------lhpevpeplrgtylgltcepvlehlkqlgvttiqllpvhakvhdrhlv 231

Qy 264 SSSKMNFWGYSTINFFSPMTRY-TSGGIKNCGRDAINEFKTFVREAHKRGIEVILDVVFN 322
Db 232 erglrnywgynplcyfapepeyatngpi-----savrefkmmvralhaagfevivdvvyn 286

Qy 323 HTAEGNENGPILSFKGVDNTTYYMLAPKGEFY--NYSGCGNTFNCNHPVVRQFIVDCLRY 380
Db 287 htgeggvlgptlsfrgidnrayykadpnnprflvdytgtgntldvgnpyviqlimdslry 346

Qy 381 WVTEMHVDGFRFDLASIMTRGSSLWDPVNVYGAPIEGDMITTGTPLVTPPLIDMISNDPI 440
Db 347 wvtemhvdgfrfdlaaalar--elydv----------dmlst--------ffqviqqdpv 386

Qy 441 LGGVKLIAEAWDAG-GLYQVGQFPHWNVWSEWNGKYRDIVRQFIKGTDGFAGGFAECLCG 499
Db 387 lsqvkliaepwdvgpggyqvghfp-wq-wtewngryrdavrrfwrgdrglngefatrfag 444

Qy 500 SPHLYQAGGRKPWHSINFVCAHDGFTLADLVTYNKKYNLPNGENNRDGENHNLSWNCGEE 559
Db 445 ssdlyersgrrpfasinfvtahdgftledlvsytkkhneanlegnrdgmdenystncgve 504

Qy 560 GEFARLSVKRLRKRQMRNFFVCLMVSQGVPMFYMGDEYGHTKGGNNNTYCHDSYVNYFRW 619
Db 505 gptqdpsvlacrealkrslistlflsqgvpmllggdelsrtqhgnnnaycqdneiswynw 564

Qy 620 D---KKEQYSELHRFCCLMTKFRKECEGLGLEDFPTAKRLQWHGHQP--------GKP-- 666
Db 565 qldtrkqqfle---fvrqtiwfrkqhrsfrrrhfltglp---nggrprrslvapegrpmr 618

Qy 667 --DWSENSRFVAFSM---------KDE-----RQGEIYVAFNTSHLPAVVELPERAG--- 707
Db 619 hedwt-npeltafglllhgdaiqgtdehgrpfrddtflilfnngseavpvvvpevcscgk 677

Qy 708 -RRWEPV------VDTGKPAPYDFLTDDLPDRALTI 736
Db 678 phhwevvpvfqrnvepptcapgetls--lppgvltv 711

RESULT 13
AAY78515
ID AAY78515 standard; protein; 726 AA.
XX
AC AAY78515;
XX
DT 05-MAY-2000 (first entry)
XX
DE Isoamylase amino acid sequence.
XX
KW Isoamylase; debranching enzyme; amylopectin; thermostable; corn;
KW genetically engineered variant; wheat; potato; wheat; manioc; rice;
KW starch; sugar production; high fructose syrup; high maltose syrup;
KW maltodextrin.
XX
OS Rhodothermus marinus.
XX
PN WO200001796-A2.
XX
PD 13-JAN-2000.
XX
PF 02-JUL-1999; 99WO-DK00381.
XX
PR 02-JUL-1998; 98DK-0000868.
XX
PA (NOVO) NOVO-NORDISK AS.
XX
PI Bisgard-Frantzen H, Svendsen A;
XX
DR WPI; 2000-160767/14.
DR N-PSDB; AAZ89961.
XX
PT Variant bacterial pullulanases and isoamylases having, e.g. increased
PT thermostability, used for converting starch from potatoes into high
PT fructose syrup -
XX
PS Claim 29; Fig 3; 116pp; English.
XX
CC This sequence represents an isoamylase amino acid sequence. Isoamylases
CC are debranching enzymes which hydrolyse alpha-1,6-D-glucosidic branch
CC linkages in amylopectin. This sequence represents a parent enzyme from
CC which genetically engineered variant debranching enzymes can be created.
CC The genetically engineered enzyme has an improved thermostability at a pH
CC in the range of 4-6 compared to the parent enzyme. The modified enzyme
CC may alternatively or also have an increased activity towards amylopectin
CC and or glycogen compared to the parent enzyme.The enzyme variants are
CC useful for converting starch to one or more sugars. This is useful when
CC starches such as corn, potato, wheat, manioc and rice starch are used as
CC starting materials in commercial large scale production of sugars, such
CC as high fructose syrup, high maltose syrup, maltodextrins, amylose, G4-G6
CC oligosaccharides and other carbohydrates such as fat replacers. The
CC thermostable debranching enzymes of the invention make it possible to
CC perform liquefaction and debranching simultaneously before the
CC saccharification step. By debranching using the thermostable debranching
CC enzyme variants during liquefaction together with the action of an
CC alpha-amylase, the formation of panose precursors is reduced, thereby
CC reducing the panose content in the final product and increasing the
CC overall saccharification yield.
XX
SQ Sequence 726 AA;

Query Match 28.2%; Score 1187; DB 21; Length 726;
Best Local Similarity 38.9%; Pred. No. 2.2e-96;
Matches 294; Conservative 86; Mismatches 242; Indels 134; Gaps 30;

Qy 57 VLAGMPAPLGATALAGGVNFAVYSGGATAAALCLFT-PEDLKADRVTEEVPLDPLMNRTG 115
Db 14 vwpgrpyplgatwdglgvnfalysqhaeavelvlfdhpddpapsrtie------vtertg 67

Qy 116 NVWHVFIEGELHNMLYGYRFDGTFAPHCGHYLDVSNVVVDPYAKAVISRGEYGVPARGNN 175
Db 68 piwhvylpglrpgqlygyrvygpyrpeeghrfnpnkvlldpyakai------grplr--- 118

```
Qy    176 CWPQM---------AGMIPL------PYS--------TFDWEGDLPLRYPQKDLVIYEMH 212
Db    119 -whdslfgykigdpagdlsfseedsapyaplgavvegcfewgddrppripwedtiiyeth 177

Qy    213 LRGFTKHDSSNVEHP-------GTFIGAVSK--LDYLKELGVNCIELMPCHEFNELEYST 263
Db    178 vkgitk------lhpevpeplrgtylglтcepvlehlkqlgvttiqllpvhakvhdrhlv 231

Qy    264 SSSKMNFWGYSTINFFSPMTRY-TSGGIKNCGRDAINEFKTFVREAHKRGIEVILDVVFN 322
Db    232 erglrnywgynplcyfapepeyatngpi-----savrefkmmvralhaagfevivdvvyn 286

Qy    323 HTAEGNENGPILSFKGVDNTTYYMLAPKGEFY--NYSGCGNTFNCNHPVVRQFIVDCLRY 380
Db    287 htgeggvlgptlsfrgidnrayykadpnnprflvdytgtgntldvgnpyviqlimdslry 346

Qy    381 WVTEMHVDGFRFDLASIMTRGSSLWDPVNVYGAPIEGDMITTGTPLVTPPLIDMISNDPI 440
Db    347 wvtemhvdgfrfdlaaalar--elydv---------dmlst--------ffqviqqdpv 386

Qy    441 LGGVKLIAEAWDAG-GLYQVGQFPHWNVWSEWNGKYRDIVRQFIKGTDGFAGGFAECLCG 499
Db    387 lsqvkliaepwdvgpggyqvghfp-wq-wtewngryrdavrrfwrgdrglngefatrfag 444

Qy    500 SPHLYQAGGRKPWHSINFVCAHDGFTLADLVTYNKKYNLPNGENNRDGENHNLSWNCGEE 559
Db    445 ssdlyersgrrpfasinfvtahdgftledlvsytkkhneanlegnrdgmdenystncgve 504

Qy    560 GEFARLSVKRLRKRQMRNFFVCLMVSQGVPMFYMGDEYGHTKGGNNNTYCHDSYVNYFRW 619
Db    505 gptqdpsvlacrealkrslistlflsqgvpmllggdelsrtqhgnnnaycqdneiswynw 564

Qy    620 D---KKEQYSELHRFCCLMTKFRKECEGLGLEDFPTAKRLQWHGHQP--------GKP-- 666
Db    565 qldtrkqqfle---fvrqtiwfrkqhrsfrrrhfltglp---nggrprrslvapegrpmr 618

Qy    667 --DWSENSRFVAFSM---------KDE-----RQGEIYVAFNTSHLPAVVELPERAG--- 707
Db    619 hedwt-npeltafglllhgdaiqgtdehgrpfrddtflilfnngseavpvvvpevcscgk 677

Qy    708 -RRWEPV------VDTGKPAPYDFLTDDLPDRALTI 736
Db    678 phhwevvpvfqrnvepptcapgetls--lppgvltv 711
```

RESULT 14
AAU60903
ID AAU60903 standard; Protein; 718 AA.
XX
AC AAU60903;
XX
DT 27-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein #21799.
XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
OS Propionibacterium acnes.
XX
PN WO200181581-A2.
XX
PD 01-NOV-2001.
XX
PF 20-APR-2001; 2001WO-US12865.
XX
PR 21-APR-2000; 2000US-199047P.
PR 02-JUN-2000; 2000US-208841P.
PR 07-JUL-2000; 2000US-216747P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
DR WPI; 2001-616774/71.
DR N-PSDB; AAS59613.
XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris -
XX
PS Example 1; SEQ ID No 22098; 1069pp; English.
XX
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 718 AA;

Query Match 25.2%; Score 1062.5; DB 22; Length 718;
Best Local Similarity 35.5%; Pred. No. 2.5e-85;
Matches 254; Conservative 102; Mismatches 285; Indels 75; Gaps 20;

```
Qy     65 LGATALAGGVNFAVYSGGATAAALCLFTPEDLKADRVTEEVPLDPLMNRTGNVWHVFIEG 124
Db     26 lgstlidggcrfglwapraervelalvadd---gsqvnhdmtrd-----ehgvwtvevsd 77

Qy    125 ELHNMLYGYRFDGTFAPHCGHYLDVSNVVVDPYAKAVISRGEYGVPARG----NNCWPQM 180
Db     78 vqaggrygyrvhgpwdpdrgmrfnpaklлldpyaraitagvdyhgpimdhtpesnyepdl 137

Qy    181 ---AGMIPLPYSTFDWEGDLPL--RYPQKDLVIYEMHLRGFTKHDSSNVEH-PGTFIGAV 234
Db    138 tddatsvplavviddpgpptpiarrhdisesviyethvkgytrlhplvpehlrgtyagla 197

Qy    235 --SKLDYLKELGVNCIELMPCHEFNELEYSTSSSKMNFWGYSTINFFSPMTRYTSGGIKN 292
Db    198 ypaviehlksigvtaiellpvqqfvsepfivgrglsnywgyntlgffaphaaycsvgsmg 257

Qy    293 CGRDAINEFKTFVREAHKRGIEVILDVVFNHTAEGNENGPILSFKGVDNTTYYMLA-PKG 351
Db    258 ---tqvrefkdmvtsfheagievfldvvynhtgegghegptlsfrgidhesyyrltndhr 314

Qy    352 EFYNYSGCGNTFNCNHPVVRQFIVDCLRYWVTEMHVDGFRFDLASIMTRGSSLWDPVNVY 411
Db    315 ndydvtgcgnsvdtshpevlamvldslrywvtemgvdgfrydlattlirdks-------h 367

Qy    412 GAPIEGDMITTGTPLVTPPLIDMISNDPILGGVKLIAEAWDAGGL-YQVGQF-PHWNVWS 469
Db    368 g--------vdqnhlfkqalve----dpilnkvkhiaepwdlgpygyqvgawgph---ws 412

Qy    470 EWNGKYRDIVRQFIKGTDGFAGGFAECLCGSPHLYQAGGRKPWHSINFVCAHDGFTLADL 529
Db    413 ewndrfrnyvrdfwrgavrgveelatrlcgspdly----gqptssvnfitahdgftmrdl 468
```

```
Qy    530 VTYNKKYNLPNGENNRDGENHNLSWNCGEEGEFARLSVKRLRKRQMRNFFVCLMVSQGVP 589
          |||| |:|  |||:| || | | :|||| |||    |:  || ||:||    |::: |||
Db    469 vtynmkhnqgngednhdgsndnrawncgwegettdqsivmlrhrqvrnlvatlllatgvp 528

Qy    590 MFYMGDEYGHTKGGNNNTYCHDSYVNYFRWDKKEQYSELHRFCCLMTKFRKECEGL---- 645
          |   ||| | |: |||| || || ::::  |:  | :|::     ::|  |:|   |
Db    529 mitagdemgrtqqgnnnaycqdspiswidwedaeawsdvtdlvkiitalrrehpalrpsr 588

Qy    646 ---------GLEDFPTAKRLQW---HGHQPGKPDWSENSRFVAFSMKDERQGEIYVAFNT 693
                  : :|    | |   || :    || : ||       :      : |:
Db    589 yrhhdpvgdandsyprradlawfsghggemvnndwhdesrrtlgmytsddneafliwfha 648

Qy    694 SHLPAVVELPE-RAGRRWEPVVDTGKP------APYDFLTDDLPDRALTIHQFSHF 742
            |  : ||  | |  |  |  |  |      |    |:  || | : : | | |
Db    649 gdrpieiilpsvrwgeswhvvastalpgeipdeavpaglsvtlpsrcvvvmqaspf 704


RESULT  15
AAY00873
ID   AAY00873 standard; Protein; 458 AA.
XX
AC   AAY00873;
XX
DT   21-MAY-1999  (first entry)
XX
DE   Original S. tuberosum isoamylase clone 9 protein sequence.
XX
KW   Isoamylase clone; potato; branched polysaccharide synthesis; amylopectin;
KW   starch modification; debranching activity.
XX
OS   Solanum tuberosum.
XX
PN   WO9906575-A1.
XX
PD   11-FEB-1999.
XX
PF   30-JUL-1998;   98WO-GB02280.
XX
PR   31-JUL-1997;   97GB-0016185.
XX
PA   (PLAN-) PLANT BIOSCIENCE LTD.
XX
PI   Bustos Guillen R,  Edwards EA,  Martin CR,  Smith AM;
XX
DR   WPI; 1999-153803/13.
DR   N-PSDB; AAX27067.
XX
PT   New nucleic acid from potato encoding starch debranching enzyme with
PT   isoamylase activity - and related vectors, transformed cells and
PT   plants, proteins and antibodies, used to generate starch with
PT   altered properties, for use in foods, as thickeners etc.
XX
PS   Disclosure; Fig 15; 78pp; English.
XX
CC   This sequence represents the original isolated clone of a
CC   Solanum tuberosum isoamylase of the invention.
CC   Fragments of the isoamylase coding sequence are used to identify and
CC   clone isoamylases from other plant species, by standard hybridisation
CC   or amplification methods. Expression of the isoamylase DNA in host cells
CC   is used to produce the isoamylase which is used: (i) to raise antibodies
CC   (Ab) (for identification, isolation and localisation of isoamylases); and
CC   (ii) for synthesis of branched polysaccharides. The DNA, proteins
CC   and Ab (or related peptides) are used to alter the quality and
CC   quantity of polysaccharides in a host cells, specifically to alter
CC   branching in amylopectin. The modified starches produced are useful in
CC   human or animal foods; as biodegradable plastic; as food or paint
CC   thickener; in starch-coated films, paper and textiles; in mining
CC   explosives; in pharmaceuticals and glues. Manipulation of debranching
CC   activity in a plant allows control of starch properties, e.g. increased
CC   gel strength; formation of paste rather than gel, changed physical
CC   characteristics etc.
XX

SQ   Sequence   458 AA;

  Query Match             24.5%;  Score 1032.5;  DB 20;  Length 458;
  Best Local Similarity   47.1%;  Pred. No. 5.9e-83;
  Matches  202;  Conservative   56;  Mismatches  130;  Indels   41;  Gaps    7;

Qy    335 SFKGVDNTTYYM--LAPKGEFYNYSGCGNTFNCNHPVVRQFIVDCLRYWVTEMHVDGFRF 392
          ||:|:||  |||  |    : |::|||||||||||| | : |:: ||:|||| |||||||
Db      1 sfrgidnkvyymvdlnnnaqllnfagcgntfncnhptvmelileslrhwvteyhvdgfrf 60

Qy    393 DLASIMTRGSSLWDPVNVYGAPIEGDMITTGTPLVTPPLIDMISNDPILGGVKLIAEAWD 452
          ||||:: ||                  | |||:  |||:  || | :|   |:||| ||
Db     61 dlasvlcrg-----------------tdgtpinapplvkaiskdsvlsrckiiaepwd 101

Qy    453 AGGLYQVGQFPHWNVWSEWNGKYRDIVRQFIKGTDGFAGGFAECLCGSPHLYQAGGRKPW 512
           |||| ||:||:|: |:|||||||| :|:|||| |  | ||  : ||  ||:  |||:
Db    102 cgglylvgkfpnwdrwaewngkyrddirrfikgdagmkgnfatriagsadlyrvnkrkpy 161

Qy    513 HSINFVCAHDGFTLADLVTYNKKYNLPNGENNRDGENHNLSWNCGEEGEFARLSVKRLRK 572
          ||:||| |||||| |||:|| |:|  |||   || | | ||||| ||| :  ::  ||
Db    162 hsvnfviahdgftlydlvsynnkhndangeggndgcndnfswncgiegetsdaninalrs 221

Qy    573 RQMRNFFVCLMVSQGVPMFYMGDEYGHTKGGNNNTYCHDSYVNYFRWDKKEQYSELH-RF 631
          |||:|| : |||||| ||  ||||||||: ||||:| ||: :| |:| : |     | ||
Db    222 rqmknfhlalmvsqgtpmmlmgdeyghtrygnnnsyghdtainnfqwgqlearkndhfrf 281

Qy    632 CCLMTKFRKECEGLGLEDFPTAKRLQWHGHQPGKPDW-SENSRFVAFSMKDERQGEIYVA 690
            | |||     |  |:|     : |      : :| :| |||:|| : |   |:||:|
Db    282 fskmikfrlshnvlrkenfieknditwl-----ednwyneesrflafmlhdgnggdiyla 336

Qy    691 FNTSHLPAVVELPE-RAGRRWEPVVDTGKPAPYDFLTDDLPDRALTIHQFSHFLYSNLYP 749
          ||  |      :|     | |  ||||   :| ||: : :                |  |
Db    337 fnahhfsiktaipspprnrswyrvvdtnlkspddfviegvsg------------isetyd 384

Qy    750 MLSYSSVIL 758
          :  ||:::|
Db    385 vapysaill 393


Search completed: July 31, 2002, 12:06:11
Job time: 4562 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: July 31, 2002, 12:00:20 ; Search time 25.38 Seconds
(without alignments)
735.270 Million cell updates/sec

Title: US-09-674-817A-2
Perfect score: 4212
Sequence: 1 SGPAPRLRRWRPNATAGKGV.....SNLYPMLSYSSVILVLRPDV 764

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3527.5	83.7	818	2	US-08-410-784A-2	Sequence 2, Appli
2	3527.5	83.7	818	4	US-09-346-237-11	Sequence 11, Appl
3	3465	82.3	740	2	US-08-410-784A-4	Sequence 4, Appli
4	2612	62.0	606	4	US-09-187-124-2	Sequence 2, Appli
5	1229.5	29.2	713	4	US-09-346-237-9	Sequence 9, Appli
6	1211	28.8	718	4	US-09-346-237-10	Sequence 10, Appl
7	1187	28.2	726	3	US-09-129-075-4	Sequence 4, Appli
8	1187	28.2	726	4	US-09-346-237-3	Sequence 3, Appli
9	1187	28.2	726	4	US-09-346-237-13	Sequence 13, Appl
10	825	19.6	774	4	US-09-346-237-8	Sequence 8, Appli
11	810	19.2	777	1	US-08-476-519-2	Sequence 2, Appli
12	810	19.2	777	5	PCT-US95-09323-2	Sequence 2, Appli
13	807	19.2	746	1	US-08-476-519-11	Sequence 11, Appl
14	807	19.2	746	5	PCT-US95-09323-11	Sequence 11, Appl
15	803.5	19.1	776	4	US-09-346-237-7	Sequence 7, Appli
16	802.5	19.1	776	4	US-09-346-237-4	Sequence 4, Appli
17	782.5	18.6	750	6	5457037-3	Patent No. 5457037
18	781.5	18.6	751	6	5457037-5	Patent No. 5457037
19	685.5	16.3	772	2	US-08-410-784A-5	Sequence 5, Appli
20	552.5	13.1	915	4	US-09-346-237-2	Sequence 2, Appli
21	544.5	12.9	820	4	US-09-313-677-21	Sequence 21, Appl
22	544.5	12.9	926	4	US-09-313-677-2	Sequence 2, Appli
23	544.5	12.9	928	4	US-09-514-599-4	Sequence 4, Appli
24	544.5	12.9	933	4	US-09-313-677-19	Sequence 19, Appl
25	544.5	12.9	967	4	US-09-313-677-17	Sequence 17, Appl
26	516	12.3	921	4	US-09-514-599-2	Sequence 2, Appli
27	497.5	11.8	928	1	US-08-474-140-11	Sequence 11, Appl
28	497.5	11.8	928	1	US-08-477-630-11	Sequence 11, Appl
29	497.5	11.8	928	1	US-08-472-293-11	Sequence 11, Appl
30	497.5	11.8	928	1	US-08-474-545-11	Sequence 11, Appl
31	497.5	11.8	928	2	US-08-478-341-11	Sequence 11, Appl
32	497.5	11.8	928	3	US-08-996-733-11	Sequence 11, Appl
33	478	11.3	829	4	US-09-514-599-6	Sequence 6, Appli
34	478	11.3	862	4	US-09-346-237-1	Sequence 1, Appli
35	449.5	10.7	893	4	US-09-514-302-4	Sequence 4, Appli
36	449.5	10.7	1938	4	US-09-514-302-2	Sequence 2, Appli
37	344	8.2	964	3	US-08-860-339-18	Sequence 18, Appl
38	339	8.0	559	4	US-09-242-690A-15	Sequence 15, Appl
39	310	7.4	597	1	US-08-399-646-4	Sequence 4, Appli
40	310	7.4	597	1	US-08-607-321-4	Sequence 4, Appli
41	310	7.4	597	2	US-08-961-240-4	Sequence 4, Appli
42	310	7.4	597	2	US-08-605-501-4	Sequence 4, Appli
43	310	7.4	598	1	US-08-399-646-14	Sequence 14, Appl
44	310	7.4	598	1	US-08-607-321-14	Sequence 14, Appl
45	310	7.4	598	2	US-08-961-240-14	Sequence 14, Appl

ALIGNMENTS

```
RESULT 1
US-08-410-784A-2
; Sequence 2, Application US/08410784A
; Patent No. 5912413
;  GENERAL INFORMATION:
;    APPLICANT:  MYERS, ALAN M.
;    APPLICANT:  JAMES, MARTHA G.
;    TITLE OF INVENTION:  ISOLATION OF SU1, A STARCH DEBRANCHING
;    TITLE OF INVENTION:  ENZYME, THE PRODUCT OF THE MAIZE GENE
;    TITLE OF INVENTION:  SUGARY 1
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Weingarten, Schurgin, Gagnebin and Hayes LLP
;      STREET:  Ten Post Office Square
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/410,784A
;      FILING DATE:  24-MAR-1995
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:.
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Heine, Ph.D., Holliday C
;      REGISTRATION NUMBER:  34,346
;      REFERENCE/DOCKET NUMBER:  ISU-002XX
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-542-2290
;      TELEFAX:  617-451-0313
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  818 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:  internal
;    ORIGINAL SOURCE:
US-08-410-784A-2
```

Query Match 83.7%; Score 3527.5; DB 2; Length 818;
Best Local Similarity 83.0%; Pred. No. 0;
Matches 649; Conservative 44; Mismatches 70; Indels 19; Gaps 6;

```
Qy       1 SGPAPRLR----RW------RPNATAGKGVGEV------CAAVVEAATKVEDEGEEDEPV 44
Db      38 SSPRPLLAVPAGRWRAGVRGRPN-VAGLGRGRLSLHAAAARPVAEAVQAEEDDDDDDEEV 96

Qy      45 AEDRYALGGACRVLAGMPAPLGATALAGGVNFAVYSGGATAAALCLFTPEDLKADRVTEE 104
Db      97 AEERFALGGACRVLAGMPAPLGATALRGGVNFAVYSSGASAASLSLFAPGDLKADRVTEE 156

Qy     105 VPLDPLMNRTGNVWHVFIEG-ELHNMLYGYRFDGTFAPHCGHYLDVSNVVVDPYAKAVIS 163
Db     157 VPLDPLLNRTGNVWHVFIHGDELHGMLCGYRFDGVFAPERGQYYDVSNVVVDPYAKAVVS 216

Qy     164 RGEYGVPARGNNCWPQMAGMIPLPYSTFDWEGDLPLRYPQKDLVIYEMHLRGFTKHDSSN 223
Db     217 RGEYGVPAPGGSCWPQMAGMIPLPYNKFDWQGDLPLGYHQKDLVIYEMHLRGFTKHNSSK 276

Qy     224 VEHPGTFIGAVSKLDYLKELGVNCIELMPCHEFNELEYSTSSSKMNFWGYSTINFFSPMT 283
Db     277 TKHPGTYIGAVSKLDHLKELGVNCIELMPCHEFNELEYFSSSSKMNFWGYSTINFFSPMA 336

Qy     284 RYTSGGIKNCGRDAINEFKTFVREAHKRGIEVILDVVFNHTAEGNENGPILSFKGVDNTT 343
Db     337 RYSSSGIRDSGCGAINEFKAFVREAHKRGIEVIMDVVFNHTAEGNEKGPILSFRGIDNST 396

Qy     344 YYMLAPKGEFYNYSGCGNTFNCNHPVVRQFIVDCLRYWVTEMHVDGFRFDLASIMTRGSS 403
Db     397 YYMLAPKGEFYNYSGCGNTFNCNHPVVREFIVDCLRYWVTEMHVDGFRFDLASILTRGCS 456

Qy     404 LWDPVNVYGAPIEGDMITTGTPLVTPPLIDMISNDPILGGVKLIAEAWDAGGLYQVGQFP 463
Db     457 LWDPVNVYGSPMEGDMITTGTPLVAPPLIDMISNDPILGNVKLIAEAWDAGGLYQEGQFP 516

Qy     464 HWNVWSEWNGKYRDIVRQFIKGTDGFAGGFAECLCGSPHLYQAGGRKPWHSINFVCAHDG 523
Db     517 HWNVWSEWNGKYRDTVRQFIKGTDGFAGAFAECLCGSPQLYQAGGRKPWHSIGFVCAHDG 576

Qy     524 FTLADLVTYNKKYNLPNGENNRDGENHNLSWNCGEEGEFARLSVKRLRKRQMRNFFVCLM 583
Db     577 FTLADLVTYNSKYNLSNGEDFRDGENHNLSWNCGEEGEFASLSVRRLRKRQMRNFFVCLM 636

Qy     584 VSQGVPMFYMGDEYGHTKGGNNNTYCHDSYVNYFRWDKK-EQYSELHRFCCLMTKFRKEC 642
Db     637 VSQGVPMFYMGDEYGHTKGGNNNTYCHDHYVNYFRWDKKEEQSSDLYRFCRLMTEFRKEC 696

Qy     643 EGLGLEDFPTAKRLQWHGHQPGKPDWSENSRFVAFSMKDERQGEIYVAFNTSHLPAVVEL 702
Db     697 ESLGLEDFPTSERLKWHGHQPGKPDWSEASRFVAFTMKDETKGEIYVAFNTSHLPVVVGL 756

Qy     703 PERAGRRWEPVVDTGKPAPYDFLTDDLPDRALTIHQFSHFLYSNLYPMLSYSSVILVLRP 762
Db     757 PERSGFRWEPVVDTGKEAPYDFLTDGLPDRAVTVYQFSHFLNSNLYPMLSYSSIILVLRP 816

Qy     763 DV 764
Db     817 DV 818
```

RESULT 2
US-09-346-237-11
; Sequence 11, Application US/09346237A
; Patent No. 6265197
; GENERAL INFORMATION:
; APPLICANT: Bisgard-Frantzen, Henrik
; APPLICANT: Svendsen, Allan
; TITLE OF INVENTION: Starch Debranching Enzymes
; FILE REFERENCE: 5629.200-US
; CURRENT APPLICATION NUMBER: US/09/346,237A
; CURRENT FILING DATE: 1999-07-01
; EARLIER APPLICATION NUMBER: PA 1998 00868
; EARLIER FILING DATE: 1998-07-02
; EARLIER APPLICATION NUMBER: 60/094,353
; EARLIER FILING DATE: 1998-07-28
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 818
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(818)
; OTHER INFORMATION: Isoamylase
US-09-346-237-11

Query Match 83.7%; Score 3527.5; DB 4; Length 818;
Best Local Similarity 83.0%; Pred. No. 0;
Matches 649; Conservative 44; Mismatches 70; Indels 19; Gaps 6;

Qy 1 SGPAPRLR----RW------RPNATAGKGVGEV------CAAVVEAATKVEDEGEEDEPV 44
Db 38 SSPRPLLAVPAGRWRAGVRGRPN-VAGLGRGRLSLHAAAARPVAEAVQAEEDDDDDDEEV 96

Qy 45 AEDRYALGGACRVLAGMPAPLGATALAGGVNFAVYSGGATAAALCLFTPEDLKADRVTEE 104
Db 97 AEERFALGGACRVLAGMPAPLGATALRGGVNFAVYSSGASAASLSLFAPGDLKADRVTEE 156

Qy 105 VPLDPLMNRTGNVWHVFIEG-ELHNMLYGYRFDGTFAPHCGHYLDVSNVVVDPYAKAVIS 163
Db 157 VPLDPLLNRTGNVWHVFIHGDELHGMLCGYRFDGVFAPERGQYYDVSNVVVDPYAKAVVS 216

Qy 164 RGEYGVPARGNNCWPQMAGMIPLPYSTFDWEGDLPLRYPQKDLVIYEMHLRGFTKHDSSN 223
Db 217 RGEYGVPAPGGSCWPQMAGMIPLPYNKFDWQGDLPLGYHQKDLVIYEMHLRGFTKHNSSK 276

Qy 224 VEHPGTFIGAVSKLDYLKELGVNCIELMPCHEFNELEYSTSSSKMNFWGYSTINFFSPMT 283
Db 277 TKHPGTYIGAVSKLDHLKELGVNCIELMPCHEFNELEYFSSSSKMNFWGYSTINFFSPMA 336

Qy 284 RYTSGGIKNCGRDAINEFKTFVREAHKRGIEVILDVVFNHTAEGNENGPILSFKGVDNTT 343
Db 337 RYSSSGIRDSGCGAINEFKAFVREAHKRGIEVIMDVVFNHTAEGNEKGPILSFRGIDNST 396

Qy 344 YYMLAPKGEFYNYSGCGNTFNCNHPVVRQFIVDCLRYWVTEMHVDGFRFDLASIMTRGSS 403
Db 397 YYMLAPKGEFYNYSGCGNTFNCNHPVVREFIVDCLRYWVTEMHVDGFRFDLASILTRGCS 456

Qy 404 LWDPVNVYGAPIEGDMITTGTPLVTPPLIDMISNDPILGGVKLIAEAWDAGGLYQVGQFP 463
Db 457 LWDPVNVYGSPMEGDMITTGTPLVAPPLIDMISNDPILGNVKLIAEAWDAGGLYQEGQFP 516

Qy 464 HWNVWSEWNGKYRDIVRQFIKGTDGFAGGFAECLCGSPHLYQAGGRKPWHSINFVCAHDG 523
Db 517 HWNVWSEWNGKYRDTVRQFIKGTDGFAGAFAECLCGSPQLYQAGGRKPWHSIGFVCAHDG 576

Qy 524 FTLADLVTYNKKYNLPNGENNRDGENHNLSWNCGEEGEFARLSVKRLRKRQMRNFFVCLM 583
Db 577 FTLADLVTYNSKYNLSNGEDFRDGENHNLSWNCGEEGEFASLSVRRLRKRQMRNFFVCLM 636

Qy 584 VSQGVPMFYMGDEYGHTKGGNNNTYCHDSYVNYFRWDKK-EQYSELHRFCCLMTKFRKEC 642
Db 637 VSQGVPMFYMGDEYGHTKGGNNNTYCHDHYVNYFRWDKKEEQSSDLYRFCRLMTEFRKEC 696

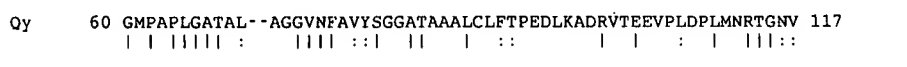
Qy 643 EGLGLEDFPTAKRLQWHGHQPGKPDWSENSRFVAFSMKDERQGEIYVAFNTSHLPAVVEL 702
Db 697 ESLGLEDFPTSERLKWHGHQPGKPDWSEASRFVAFTMKDETKGEIYVAFNTSHLPVVVGL 756

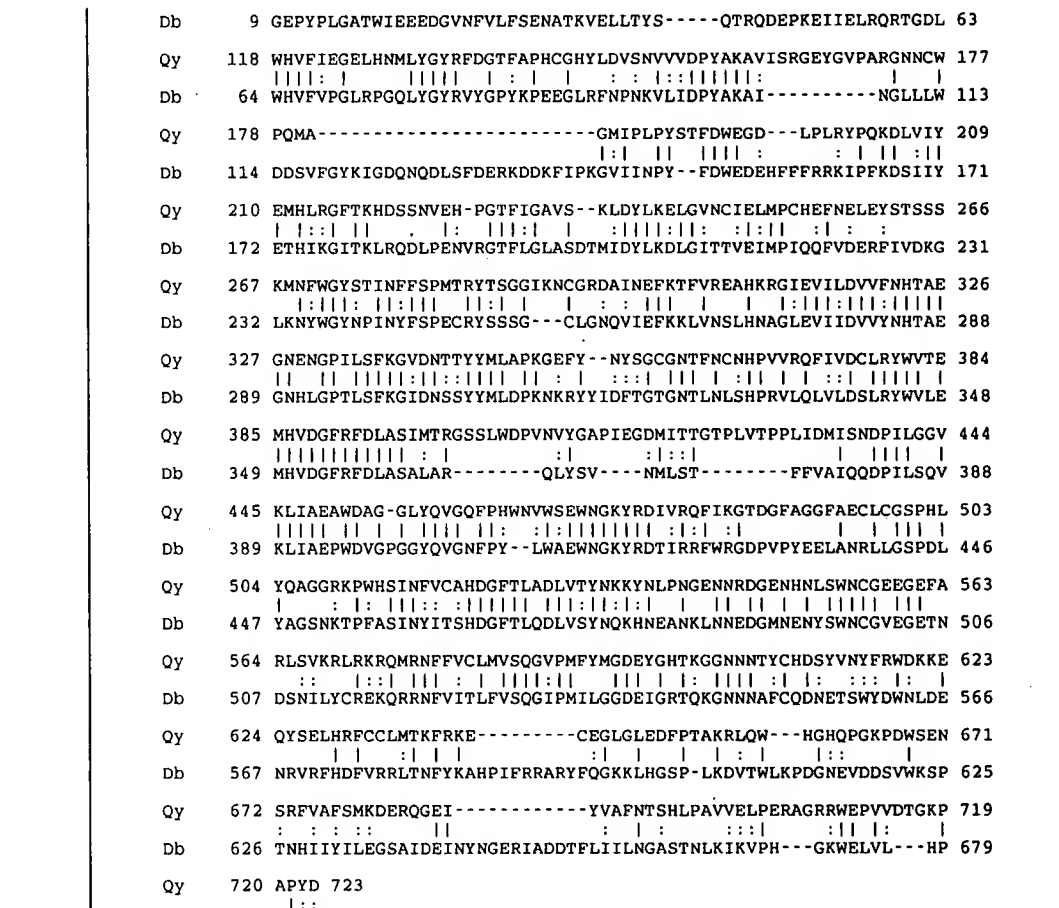
Qy 703 PERAGRRWEPVVDTGKPAPYDFLTDDLPDRALTIHQFSHFLYSNLYPMLSYSSVILVLRP 762
Db 757 PERSGFRWEPVVDTGKEAPYDFLTDGLPDRAVTVYQFSHFLNSNLYPMLSYSSIILVLRP 816

```
Qy     763 DV 764
           ||
Db     817 DV 818


RESULT   3
US-08-410-784A-4
; Sequence 4, Application US/08410784A
; Patent No. 5912413
;  GENERAL INFORMATION:
;    APPLICANT:  MYERS, ALAN M.
;    APPLICANT:  JAMES, MARTHA G.
;    TITLE OF INVENTION:  ISOLATION OF SU1, A STARCH DEBRANCHING
;    TITLE OF INVENTION:  ENZYME, THE PRODUCT OF THE MAIZE GENE
;    TITLE OF INVENTION:  SUGARY 1
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Weingarten, Schurgin, Gagnebin and Hayes LLP
;      STREET:  Ten Post Office Square
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/410,784A
;      FILING DATE:  24-MAR-1995
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Heine, Ph.D., Holliday C
;      REGISTRATION NUMBER:  34,346
;      REFERENCE/DOCKET NUMBER:  ISU-002XX
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-542-2290
;      TELEFAX:  617-451-0313
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  740 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:  N-terminal
;    ORIGINAL SOURCE:
US-08-410-784A-4


  Query Match             82.3%;  Score 3465;  DB 2;  Length 740;
  Best Local Similarity   85.3%;  Pred. No. 0;
  Matches  631;  Conservative   43;  Mismatches   62;  Indels    4;  Gaps    3;

Qy      27 VVEAATKVEDEGEEDEPVAEDRYALGGACRVLAGMPAPLGATALAGGVNFAVYSGGATAA 86
           | ||    ||: ::|| |||:|:||||||||||||||||||| |||||||| ||:||
Db       3 VAEAVQAEEDDDDDDEEVAEERFALGGACRVLAGMPAPLGATALRGGVNFAVYSSGASAA 62

Qy      87 ALCLFTPEDLKADRVTEEVPLDPLMNRTGNVWHVFIEG-ELHNMLYGYRFDGTFAPHCGH 145
           :| || | ||||||||||||||:||||||||||| | ||| || ||||||  ||  |
Db      63 SLSLFAPGDLKADRVTEEVPLDPLLNRTGNVWHVFIHGDELHGMLCGYRFDGVRAPERGQ 122

Qy     146 YLDVSNVVVDPYAKAVISRGEYGVPARGNNCWPQMAGMIPLPYSTFDWEGDLPLRYPQKD 205
           | |||||||||||||:||||||||| | :|||||  ||||||: |||:||||| | |||
Db     123 YYDVSNVVVDPYAKAVVSRGEYGVPAPGGSCWPQM--MIPLPYNKFDWQGDLPLGYHQKD 180

Qy     206 LVIYEMHLRGFTKHDSSNVEHPGTFIGAVSKLDYLKELGVNCIELMPCHEFNELEYSTSS 265
           |||||||||||||:|| :||||:||||||||:|||||||||||||||||||||| :||
Db     181 LVIYEMHLRGFTKHNSSKTKHPGTYIGAVSKLDHLKELGVNCIELMPCHEFNELEYFSSS 240

Qy     266 SKMNFWGYSTINFFSPMTRYTSGGIKNCGRDAINEFKTFVREAHKRGIEVILDVVFNHTA 325
           |||||||||||||||| ||:| ||:: |  |||||| |||||||||||||:||||||||
Db     241 SKMNFWGYSTINFFSPMARYSSSGIRDSGCGAINEFKAFVREAHKRGIEVIMDVVFNHTA 300

Qy     326 EGNENGPILSFKGVDNTTYYMLAPKGEFYNYSGCGNTFNCNHPVVRQFIVDCLRYWVTEM 385
           |||| ||||||:|:||:|||||||||||||||||||||||||||||:||||||||||||
Db     301 EGNEKGPILSFRGIDNSTYYMLAPKGEFYNYSGCGNTFNCNHPVVREFIVDCLRYWVTEM 360

Qy     386 HVDGFRFDLASIMTRGSSLWDPVNVYGAPIEGDMITTGTPLVTPPLIDMISNDPILGGVK 445
           |||||||||||:||| |||||||||| |:|||||||||||| ||||||||||||| ||
Db     361 HVDGFRFDLASILTRGCSLWDPVNVYGEPMEGDMITTGTPLVAPPLIDMISNDPILGNVK 420

Qy     446 LIAEAWDAGGLYQVGQFPHWNVWSEWNGKYRDIVRQFIKGTDGFAGGFAECLCGSPHLYQ 505
           |||||||||||| ||||||:|||||||||| |||||||||||| |||||||||| |||
Db     421 LIAEAWDAGGLYQEGQFPHWHVWSEWNGKYRDTVRQFIKGTDGFAGAFAECLCGSPQLYQ 480

Qy     506 AGGRKPWHSINFVCAHDGFTLADLVTYNKKYNLPNGENNRDGENHNLSWNCGEEGEFARL 565
           |||||||||| |||||||||||||||| |||| |||: |||||||||||||||||| |
Db     481 AGGRKPWHSIGFVCAHDGFTLADLVTYNSKYNLSNGEDFRDGENHNLSWNCGEEGEFASL 540

Qy     566 SVKRLRKRQMRNFFVCLMVSQGVPMFYMGDEYGHTKGGNNNTYCHDSYVNYFRWDKK-EQ 624
           ||:|||||||||||||||||||||||||||||||||||||||||| |||||||| ||
Db     541 SVRRLRKRQMRNFFVCLMVSQGVPMFYMGDEYGHTKGGNNNTYCHDHYVNYFRWDKKEEQ 600

Qy     625 YSELHRFCCLMTKFRKECEGLGLEDFPTAKRLQWHGHQPGKPDWSENSRFVAFSMKDERQ 684
           |:|:||| |||:|||||| ||||||||::||:|||||||||||| ||||||:|||| :
Db     601 SSDLYRFCRLMTEFRKECESLGLEDFPTSERLKWHGHQPGKPDWSEASRFVAFTMKDETK 660

Qy     685 GEIYVAFNTSHLPAVVELPERAGRRWEPVVDTGKPAPYDFLTDDLPDRALTIHQFSHFLY 744
           |||||||||||| || ||||:| |||||||||| |||||||| |||||:|::|||||
Db     661 GEIYVAFNTSHLPVVVGLPERSGFRWEPVVDTGKEAPYDFLTDGLPDRAVTVYQFSHFLN 720

Qy     745 SNLYPMLSYSSVILVLRPDV 764
           ||||||||||:|||| |||
Db     721 SNLYPMLSYSSIILVLTPDV 740


RESULT   4
US-09-187-124-2
; Sequence 2, Application US/09187124A
; Patent No. 6255563
; GENERAL INFORMATION:
;  APPLICANT: Emmermann, Michael
;  APPLICANT: Kossmann, Jens
;  TITLE OF INVENTION: NUCLEIC ACID MOLECULES CODING FOR DEBRANCHING ENZYMES
;  TITLE OF INVENTION: FROM POTATO
;  FILE REFERENCE: GFB8
;  CURRENT APPLICATION NUMBER: US/09/187,124A
;  CURRENT FILING DATE: 1998-11-05
;  EARLIER APPLICATION NUMBER: PCT/EP97/02292
;  EARLIER FILING DATE: 1997-05-06
;  EARLIER APPLICATION NUMBER: DE 196 18 125.9
;  EARLIER FILING DATE: 1996-05-06
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 606
;   TYPE: PRT
;   ORGANISM: Solanum tuberosum
US-09-187-124-2


  Query Match             62.0%;  Score 2612;  DB 4;  Length 606;
  Best Local Similarity   76.6%;  Pred. No. 2.4e-246;
  Matches  458;  Conservative   62;  Mismatches   70;  Indels    8;  Gaps    3;

Qy     171 ARG--NNCWPQMAGMIPLPYSTFDWEGDLPLRYPQKDLVIYEMHLRGFTKHDSSNVEHPG 228
           |||  ::||| ||||:|     ||||||| |::||:||||||||:|||| |:||  ::||
```

RESULT 5
US-09-346-237-9
; Sequence 9, Application US/09346237A
; Patent No. 6265197
; GENERAL INFORMATION:
; APPLICANT: Bisgard-Frantzen, Henrik
; APPLICANT: Svendsen, Allan
; TITLE OF INVENTION: Starch Debranching Enzymes
; FILE REFERENCE: 5629.200-US
; CURRENT APPLICATION NUMBER: US/09/346,237A
; CURRENT FILING DATE: 1999-07-01
; EARLIER APPLICATION NUMBER: PA 1998 00868
; EARLIER FILING DATE: 1998-07-02
; EARLIER APPLICATION NUMBER: 60/094,353
; EARLIER FILING DATE: 1998-07-28
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 9
; LENGTH: 713
; TYPE: PRT
; ORGANISM: Sulfolobus acidocaldarius
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(713)
; OTHER INFORMATION: Isoamylase
US-09-346-237-9

Query Match 29.2%; Score 1229.5; DB 4; Length 713;
Best Local Similarity 38.1%; Pred. No. 3.9e-111;
Matches 276; Conservative 101; Mismatches 238; Indels 109; Gaps 21;





RESULT 6
US-09-346-237-10
; Sequence 10, Application US/09346237A
; Patent No. 6265197
; GENERAL INFORMATION:
; APPLICANT: Bisgard-Frantzen, Henrik
; APPLICANT: Svendsen, Allan
; TITLE OF INVENTION: Starch Debranching Enzymes
; FILE REFERENCE: 5629.200-US
; CURRENT APPLICATION NUMBER: US/09/346,237A
; CURRENT FILING DATE: 1999-07-01
; EARLIER APPLICATION NUMBER: PA 1998 00868
; EARLIER FILING DATE: 1998-07-02
; EARLIER APPLICATION NUMBER: 60/094,353
; EARLIER FILING DATE: 1998-07-28
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 718
; TYPE: PRT
; ORGANISM: Sulfolobus sulfataricus
; FEATURE:

```
;   NAME/KEY: PEPTIDE
;   LOCATION: (1)..(718)
;   OTHER INFORMATION: Isoamylase
US-09-346-237-10


  Query Match             28.8%;  Score 1211;  DB 4;  Length 718;
  Best Local Similarity   38.2%;  Pred. No. 2.5e-109;
  Matches  274;  Conservative  101;  Mismatches  236;  Indels  106;  Gaps   19;

Qy      60 GMPAPLGATALA--GGVNFAVYSGGATAAALCLFTPEDLKADRVTEEVPLDPLMNRTGNV 117
           | | |||:  :    ||||:::|  |    | |::  : |  :   ||      |:||::
Db      15 GDPYPLGSNWIEDDDGVNFSLFSENAEKVELLLYSLTNQKYPKEIIEV-----KNKTGDI 69

Qy     118 WHVFIEGELHNMLYGYRFDGTFAPHCGHYLDVSNVVVDPYAKAV---------------- 161
           ||||: |     || ||  | : |  |   : : |::||||||:
Db      70 WHVFVPGLRPGQLYAYRVYGPYKPELGLRFNPNKVLIDPYAKAINGSVIWNDAVFGYKIG 129

Qy     162 ----------ISRGEYGVPARGNNCWPQMAGMIPLPYSTFDWEGD---LPLRYPQKDLVI 208
                     ||| ||            ::  ||  |:|: :      : | || ||
Db     130 DQNQDLTYDERDSGEY-VP----------KSVVINPY--FEWDDEDFIKGKKVPLKDTVI 176

Qy     209 YEMHLRGFTKHDSSNVEH-PGTFIGAVSK--LDYLKELGVNCIELMPCHEFNELEYSTSS 265
           ||:|::||||       |:  ||: |  |:  : |||:||:  :|||||   | :  : |
Db     177 YEVHVKGFTKLRLDLPENIRGTYEGLASEQMISYLKDLGITTVELMPVFHFIDQRFLTDK 236

Qy     266 SKMNFWGYSTINFFSPMTRYTSGGIKNCGRDAINEFKTFVREAHKRGIEVILDVVFNHTA 325
             |:|||  ||||||  ||:| |   |     :  ||  | | |  |||||:|||:||||
Db     237 GLTNYWGYDPINFFSPECRYSSTG---CLGGQVLSFKKMVNELHNAGIEVIIDVVYNHTA 293

Qy     326 EGNENGPILSFKGVDNTTYYMLAPKGEFY--NYSGCGNTFNCNHPVVRQFIVDCLRYWVT 383
           |||  || |||:|:||| |||| |  : |  :::| ||| | :|| | | ::| ||||||
Db     294 EGNHLGPTLSFRGIDNTAYYMLQPDNKRYYLDFTGTGNTLNLSHPRVIQMVLDSLRYWVT 353

Qy     384 EMHVDGFRFDLASIMTRGSSLWDPVNVYGAPIEGDMITTGTPLVTPPLIDMISNDPILGG 443
           ||||||||||||: : |         :|      :|: |           :  ||||
Db     354 EMHVDGFRFDLAAALAR--------ELYSV----NMLNT--------FFIALQQDPILSQ 393

Qy     444 VKLIAEAWDAG-GLYQVGQFPHWNVWSEWNGKYRDIVRQFIKGTDGFAGGFAECLCGSPH 502
           |||||| || | | |||| ||:   |:|||||||| :|:| :|         |  | |||
Db     394 VKLIAEPWDVGQGGYQVGNFPY--QWAEWNGKYRDSIRRFWRGEALPYSEIANRLLGSPD 451

Qy     503 LYQAGGRKPWHSINFVCAHDGFTLADLVTYNKKYNLPNGENNRDGENHNLSWNCGEEGEF 562
           :|    : |: |||:| :|||||| |||:||:|:|  || ||:|| | | ||||| ||
Db     452 IYLGNNKTPFASINYVTSHDGFTLEDLVSYNQKHNEANGFNNQDGMNENYSWNCGAEGPT 511

Qy     563 ARLSVKRLRKRQMRNFFVCLMVSQGVPMFYMGDEYGHTKGGNNNTYCHDSYVNYFRWDKK 622
             :|   |::| ||| : |:|||| ||   |||   |: |||| :| |: : :| |:
Db     512 NDQNVVICREKQKRNFMITLLVSQGTPMILGGDELSRTQRGNNNAFCQDNEITWFDWNLD 571

Qy     623 EQYSELHRFCCLMTKFRKECEGLGLEDFPTAKRL----------QWHGHQPGKPDWSEN 671
           |: |:   |  | :| :       | :  |:|                | :  :  ||
Db     572 ERKSKFLEFVKKMIQFYRAHPAFRRERYFQGKKLFGMPLKDVTFYTLEGREVDEKTWSSP 631

Qy     672 SRFVAFSMKDERQGEI----------YVAFNTSHLPAVVELPERAGRRWEPVVDT 716
           :: | | ::     ||           :  | :     |: |:    :|| |: :
Db     632 TQLVIFVLEGSVMDEINMYGERIADDSFLIILNANPNNVKVKFPK---GKWELVISS 685


RESULT   7
US-09-129-075-4
; Sequence 4, Application US/09129075
; Patent No. 6087149
; GENERAL INFORMATION:
;  APPLICANT: Tsutsumi, No. 60871491ko
;  APPLICANT: Bisgard-Frantzen, Henrik
;  APPLICANT: Svendsen, Allan
;  TITLE OF INVENTION: Starch Conversion Process
;  FILE REFERENCE: 5262.204-US
;  CURRENT APPLICATION NUMBER: US/09/129,075
;  CURRENT FILING DATE: 1998-08-04
;  EARLIER APPLICATION NUMBER: 0787/97
;  EARLIER FILING DATE: 1997-07-02
;  EARLIER APPLICATION NUMBER: 60/055,867
;  EARLIER FILING DATE: 1997-08-13
;  EARLIER APPLICATION NUMBER: PCT/DK98/00304
;  EARLIER FILING DATE: 1998-07-02
;  NUMBER OF SEQ ID NOS: 16
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 4
;   LENGTH: 726
;   TYPE: PRT
;   ORGANISM: Rhodothermus marinus
US-09-129-075-4


  Query Match             28.2%;  Score 1187;  DB 3;  Length 726;
  Best Local Similarity   38.9%;  Pred. No. 5.7e-107;
  Matches  294;  Conservative   86;  Mismatches  242;  Indels  134;  Gaps   30;

Qy      57 VLAGMPAPLGATALAGGVNFAVYSGGATAAALCLFT-PEDLKADRVTEEVPLDPLMNRTG 115
           |  | | |||||    |||||:||  | |  | ||  |:|    |  |      :  |||
Db      14 VWPGRPYPLGATWDGLGVNFALYSQHAEAVELVLFDHPDDPAPSRTIE------VTERTG 67

Qy     116 NVWHVFIEGELHNMLYGYRFDGTFAPHCGHYLDVSNVVVDPYAKAVISRGEYGVPARGNN 175
           :|||:: |     |||||  | : |  ||  : : |::||||||:      | | |
Db      68 PIWHVYLPGLRPGQLYGYRVYGPYRPEEGHRFNPNKVLLDPYAKAI------GRPLR--- 118

Qy     176 CWPQM---------AGMIPL------PYS--------TFDWEGDLPLRYPQKDLVIYEMH 212
           |             || :        ||:         |:|  | | | | :| :||| |
Db     119 -WHDSLFGYKIGDPAGDLSFSEEDSAPYAPLGAVVEGCFEWGDDRPPRIPWEDTIIYETH 177

Qy     213 LRGFTKHDSSNVEHP-------GTFIGAVSK--LDYLKELGVNCIELMPCHEFNELEYST 263
           ::| ||       ||       ||::|   :  |::||:|||  |:|:| |      :
Db     178 VKGITK------LHPEVPEPLRGTYLGLTCEPVLEHLKQLGVTTIQLLPVHAKVHDRHLV 231

Qy     264 SSSKMNFWGYSTINFFSPMTRY-TSGGIKNCGRDAINEFKTFVREAHKRGIEVILDVVFN 322
             |:|||: : :|:|   | |:| |      |: |||  ||  |  | |||:|||:|
Db     232 ERGLRNYWGYNPLCYFAPEPEYATNGPI-----SAVREFKMMVRALHAAGFEVIVDVVYN 286

Qy     323 HTAEGNENGPILSFKGVDNTTYYMLAPKGEFY--NYSGCGNTFNCNHPVVRQFIVDCLRY 380
           || ||   || |||:|:||  ||   |    :  :|:| ||| :  :| | | |:| |||
Db     287 HTGEGGVLGPTLSFRGIDNRAYYKADPNNPRFLVDYTGTGNTLDVGNPYVIQLIMDSLRY 346

Qy     381 WVTEMHVDGFRFDLASIMTRGSSLWDPVNVYGAPIEGDMITTGTPLVTPPLIDMISNDPI 440
           |||||||||||||||: : |   |:|          ||::|                :|  ||:
Db     347 WVTEMHVDGFRFDLAAALAR--ELYDV----------DMLST--------FFQVIQQDPV 386

Qy     441 LGGVKLIAEAWDAG-GLYQVGQFPHWNVWSEWNGKYRDIVRQFIKGTDGFAGGFAECLCG 499
           |  ||||| || | | |||| || |  |:||||:||| ||:| :|  |  | ||    |
Db     387 LSQVKLIAEPWDVGPGGYQVGHFP-WQ-WTEWNGRYRDAVRRFWRGDRGLNGEFATRFAG 444

Qy     500 SPHLYQAGGRKPWHSINFVCAHDGFTLADLVTYNKKYNLPNGENNRDGENHNLSWNCGEE 559
           |  ||:  ||:|: |||| |||||| |||:| ||:|  | | |||| : | | ||| |
Db     445 SSDLYERSGRRPFASINFVTAHDGFTLEDLVSYTKKHNEANLEGNRDGMDENYSTNCGVE 504

Qy     560 GEFARLSVKRLRKRQMRNFFVCLMVSQGVPMFYMGDEYGHTKGGNNNTYCHDSYVNYFRW 619
           |    ||   |:   |:    | :||||||   |||   |: |||| || |: :::: |
Db     505 GPTQDPSVLACREALKRSLISTLFLSQGVPMLLGGDELSRTQHGNNNAYCQDNEISWYNW 564

Qy     620 D---KKEQYSELHRFCCLMTKFRKECEGLGLEDFPTAKRLQWHGHQP--------GKP-- 666
             :|:|: |   |      |||:         | |       :| :|         |:|
Db     565 QLDTRKQQFLE---FVRQTIWFRKQHRSFRRRHFLTGLP---NGGRPRRSLVAPEGRPMR 618

Qy     667 --DWSENSRFVAFSM---------KDE-----RQGEIYVAFNTSHLPAVVELPERAG--- 707
             ||: |    || :           ||      |      : ||        | :||
Db     619 HEDWT-NPELTAFGLLLHGDAIQGTDEHGRPFRDDTFLILFNNGSEAVPVVVPEVCSCGK 677

Qy     708 -RRWEPV------VDTGKPAPYDFLTDDLPDRALTI 736
              || |      |:    || : |:  ||   ||:
Db     678 PHHWEVVPVFQRNVEPPTCAPGETLS--LPPGVLTV 711


RESULT   8
US-09-346-237-3
; Sequence 3, Application US/09346237A
```

; Patent No. 6265197
; GENERAL INFORMATION:
; APPLICANT: Bisgard-Frantzen, Henrik
; APPLICANT: Svendsen, Allan
; TITLE OF INVENTION: Starch Debranching Enzymes
; FILE REFERENCE: 5629.200-US
; CURRENT APPLICATION NUMBER: US/09/346,237A
; CURRENT FILING DATE: 1999-07-01
; EARLIER APPLICATION NUMBER: PA 1998 00868
; EARLIER FILING DATE: 1998-07-02
; EARLIER APPLICATION NUMBER: 60/094,353
; EARLIER FILING DATE: 1998-07-28
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 726
; TYPE: PRT
; ORGANISM: Rhodotermus marinus
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(726)
; OTHER INFORMATION: Isoamylase
US-09-346-237-3

Query Match 28.2%; Score 1187; DB 4; Length 726;
Best Local Similarity 38.9%; Pred. No. 5.7e-107;
Matches 294; Conservative 86; Mismatches 242; Indels 134; Gaps 30;

```
Qy      57 VLAGMPAPLGATALAGGVNFAVYSGGATAAALCLFT-PEDLKADRVTEEVPLDPLMNRTG 115
           |  | | |||||    |||||:||  | |  | ||  |:|    |  |      :  |||
Db      14 VWPGRPYPLGATWDGLGVNFALYSQHAEAVELVLFDHPDDPAPSRTIE------VTERTG 67

Qy     116 NVWHVFIEGELHNMLYGYRFDGTFAPHCGHYLDVSNVVVDPYAKAVISRGEYGVPARGNN 175
           :|||:: |     |||||  | : |  ||  : : |::||||||:      | | |
Db      68 PIWHVYLPGLRPGQLYGYRVYGPYRPEEGHRFNPNKVLLDPYAKAI------GRPLR--- 118

Qy     176 CWPQM---------AGMIPL------PYS--------TFDWEGDLPLRYPQKDLVIYEMH 212
           |              || :         ||:          |:|  | | | | :| :||| |
Db     119 -WHDSLFGYKIGDPAGDLSFSEEDSAPYAPLGAVVEGCFEWGDDRPPRIPWEDTIIYETH 177

Qy     213 LRGFTKHDSSNVEHP-------GTFIGAVSK--LDYLKELGVNCIELMPCHEFNELEYST 263
           ::| ||       ||       ||::|   :  |::||:|||  |:|:| |      :
Db     178 VKGITK------LHPEVPEPLRGTYLGLTCEPVLEHLKQLGVTTIQLLPVHAKVHDRHLV 231

Qy     264 SSSKMNFWGYSTINFFSPMTRY-TSGGIKNCGRDAINEFKTFVREAHKRGIEVILDVVFN 322
           |:|||: : :|:|   | |:| |      |: |||  ||  |  | |||:|||:|
Db     232 ERGLRNYWGYNPLCYFAPEPEYATNGPI-----SAVREFKMMVRALHAAGFEVIVDVVYN 286

Qy     323 HTAEGNENGPILSFKGVDNTTYYMLAPKGEFY--NYSGCGNTFNCNHPVVRQFIVDCLRY 380
           || ||   || |||:|:||  ||   |    :  :|:| ||| :  :| | | |:| |||
Db     287 HTGEGGVLGPTLSFRGIDNRAYYKADPNNPRFLVDYTGTGNTLDVGNPYVIQLIMDSLRY 346

Qy     381 WVTEMHVDGFRFDLASIMTRGSSLWDPVNVYGAPIEGDMITTGTPLVTPPLIDMISNDPI 440
           |||||||||||||||: : |   |:|           ||::|           :|  ||:
Db     347 WVTEMHVDGFRFDLAAALAR--ELYDV----------DMLST--------FFQVIQQDPV 386

Qy     441 LGGVKLIAEAWDAG-GLYQVGQFPHWNVWSEWNGKYRDIVRQFIKGTDGFAGGFAECLCG 499
           |  |||||| || | | |||| || |  |:||||:||| ||:|  :|  |  | ||    |
Db     387 LSQVKLIAEPWDVGPGGYQVGHFP-WQ-WTEWNGRYRDAVRRFWRGDRGLNGEFATRFAG 444

Qy     500 SPHLYQAGGRKPWHSINFVCAHDGFTLADLVTYNKKYNLPNGENNRDGENHNLSWNCGEE 559
           |  ||:  ||:|: |||||  |||||||  |||:| ||:|  | | |||| : | | ||| |
Db     445 SSDLYERSGRRPFASINFVTAHDGFTLEDLVSYTKKHNEANLEGNRDGMDENYSTNCGVE 504

Qy     560 GEFARLSVKRLRKRQMRNFFVCLMVSQGVPMFYMGDEYGHTKGGNNNTYCHDSYVNYFRW 619
           |     ||   |:   |:    | :||||||   |||   |: |||| || |: :::: |
Db     505 GPTQDPSVLACREALKRSLISTLFLSQGVPMLLGGDELSRTQHGNNNAYCQDNEISWYNW 564

Qy     620 D---KKEQYSELHRFCCLMTKFRKECEGLGLEDFPTAKRLQWHGHQP--------GKP-- 666
               :|:|: |   |      |||:        | |      :| :|        |:|
Db     565 QLDTRKQQFLE---FVRQTIWFRKQHRSFRRRHFLTGLP---NGGRPRRSLVAPEGRPMR 618

Qy     667 --DWSENSRFVAFSM---------KDE-----RQGEIYVAFNTSHLPAVVELPERAG--- 707
             ||: |   || :          ||     |     : ||   .   | :||
Db     619 HEDWT-NPELTAFGLLLHGDAIQGTDEHGRPFRDDTFLILFNNGSEAVPVVVPEVCSCGK 677

Qy     708 -RRWEPV------VDTGKPAPYDFLTDDLPDRALTI 736
            || |      |:    || : |:  ||   ||:
Db     678 PHHWEVVPVFQRNVEPPTCAPGETLS--LPPGVLTV 711

RESULT 9
US-09-346-237-13
; Sequence 13, Application US/09346237A
; Patent No. 6265197
; GENERAL INFORMATION:
; APPLICANT: Bisgard-Frantzen, Henrik
; APPLICANT: Svendsen, Allan
; TITLE OF INVENTION: Starch Debranching Enzymes
; FILE REFERENCE: 5629.200-US
; CURRENT APPLICATION NUMBER: US/09/346,237A
; CURRENT FILING DATE: 1999-07-01
; EARLIER APPLICATION NUMBER: PA 1998 00868
; EARLIER FILING DATE: 1998-07-02
; EARLIER APPLICATION NUMBER: 60/094,353
; EARLIER FILING DATE: 1998-07-28
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 13
; LENGTH: 726
; TYPE: PRT
; ORGANISM: Rhodotermus marinus DSM 4252
US-09-346-237-13

Query Match 28.2%; Score 1187; DB 4; Length 726;
Best Local Similarity 38.9%; Pred. No. 5.7e-107;
Matches 294; Conservative 86; Mismatches 242; Indels 134; Gaps 30;

```
Qy   57 VLAGMPAPLGATALAGGVNFAVYSGGATAAALCLFT-PEDLKADRVTEEVPLDPLMNRTG 115
           | | | ||||| |||||:|| | | | || |:| | | : |||
Db   14 VWPGRPYPLGATWDGLGVNFALYSQHAEAVELVLFDHPDDPAPSRTIE------VTERTG 67

Qy   116 NVWHVFIEGELHNMLYGYRFDGTFAPHCGHYLDVSNVVVDPYAKAVISRGEYGVPARGNN 175
           :|||:: | ||||| | : | || : : |::||||||: | | |
Db   68 PIWHVYLPGLRPGQLYGYRVYGPYRPEEGHRFNPNKVLLDPYAKAI------GRPLR--- 118

Qy   176 CWPQM---------AGMIPL------PYS--------TFDWEGDLPLRYPQKDLVIYEMH 212
           | || : ||: |:| | | | | :| :||| |
Db   119 -WHDSLFGYKIGDPAGDLSFSEEDSAPYAPLGAVVEGCFEWGDDRPPRIPWEDTIIYETH 177

Qy   213 LRGFTKHDSSNVEHP-------GTFIGAVSK--LDYLKELGVNCIELMPCHEFNELEYST 263
           ::| || || ||::| : |::||:||| |:|:| | :
Db   178 VKGITK------LHPEVPEPLRGTYLGLTCEPVLEHLKQLGVTTIQLLPVHAKVHDRHLV 231

Qy   264 SSSKMNFWGYSTINFFSPMTRY-TSGGIKNCGRDAINEFKTFVREAHKRGIEVILDVVFN 322
           |:|||: : :|:| | |:| | |: ||| || | | |||:|||:|
Db   232 ERGLRNYWGYNPLCYFAPEPEYATNGPI-----SAVREFKMMVRALHAAGFEVIVDVVYN 286

Qy   323 HTAEGNENGPILSFKGVDNTTYYMLAPKGEFY--NYSGCGNTFNCNHPVVRQFIVDCLRY 380
           || || || |||:|:|| || | : :|:| ||| : :| | | |:| |||
Db   287 HTGEGGVLGPTLSFRGIDNRAYYKADPNNPRFLVDYTGTGNTLDVGNPYVIQLIMDSLRY 346

Qy   381 WVTEMHVDGFRFDLASIMTRGSSLWDPVNVYGAPIEGDMITTGTPLVTPPLIDMISNDPI 440
           |||||||||||||||: : | |:| ||::| :| ||:
Db   347 WVTEMHVDGFRFDLAAALAR--ELYDV----------DMLST--------FFQVIQQDPV 386

Qy   441 LGGVKLIAEAWDAG-GLYQVGQFPHWNVWSEWNGKYRDIVRQFIKGTDGFAGGFAECLCG 499
           | |||||| || | | |||| || | |:||||:||| ||:| :| | | || |
Db   387 LSQVKLIAEPWDVGPGGYQVGHFP-WQ-WTEWNGRYRDAVRRFWRGDRGLNGEFATRFAG 444

Qy   500 SPHLYQAGGRKPWHSINFVCAHDGFTLADLVTYNKKYNLPNGENNRDGENHNLSWNCGEE 559
           | ||: ||:|: ||||| ||||||| |||:| ||:| | | |||| : | | ||| |
Db   445 SSDLYERSGRRPFASINFVTAHDGFTLEDLVSYTKKHNEANLEGNRDGMDENYSTNCGVE 504
```
```

```
Qy    560 GEFARLSVKRLRKRQMRNFFVCLMVSQGVPMFYMGDEYGHTKGGNNNTYCHDSYVNYFRW 619
Db    505 GPTQDPSVLACREALKRSLISTLFLSQGVPMLLGGDELSRTQHGNNNAYCQDNEISWYNW 564

Qy    620 D---KKEQYSELHRFCCLMTKFRKECEGLGLEDFPTAKRLQWHGHQP--------GKP-- 666
Db    565 QLDTRKQQFLE---FVRQTIWFRKQHRSFRRRHFLTGLP---NGGRPRRSLVAPEGRPMR 618

Qy    667 --DWSENSRFVAFSM---------KDE-----RQGEIYVAFNTSHLPAVVELPERAG--- 707
Db    619 HEDWT-NPELTAFGLLLHGDAIQGTDEHGRPFRDDTFLILFNNGSEAVPVVVPEVCSCGK 677

Qy    708 -RRWEPV------VDTGKPAPYDFLTDDLPDRALTI 736
Db    678 PHHWEVVPVFQRNVEPPTCAPGETLS--LPPGVLTV 711


RESULT 10
US-09-346-237-8
; Sequence 8, Application US/09346237A
; Patent No. 6265197
; GENERAL INFORMATION:
;  APPLICANT: Bisgard-Frantzen, Henrik
;  APPLICANT: Svendsen, Allan
;  TITLE OF INVENTION: Starch Debranching Enzymes
;  FILE REFERENCE: 5629.200-US
;  CURRENT APPLICATION NUMBER: US/09/346,237A
;  CURRENT FILING DATE: 1999-07-01
;  EARLIER APPLICATION NUMBER: PA 1998 00868
;  EARLIER FILING DATE: 1998-07-02
;  EARLIER APPLICATION NUMBER: 60/094,353
;  EARLIER FILING DATE: 1998-07-28
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 8
;  LENGTH: 774
;  TYPE: PRT
;  ORGANISM: Favobacterium odoratum
;  FEATURE:
;  NAME/KEY: PEPTIDE
;  LOCATION: (1)..(774)
;  OTHER INFORMATION: Isoamylase
US-09-346-237-8


  Query Match           19.6%;  Score 825;  DB 4;  Length 774;
  Best Local Similarity  33.2%;  Pred. No. 1.5e-71;
  Matches  251;  Conservative   90;  Mismatches  282;  Indels  132;  Gaps   32;

Qy     50 ALGGACRVLAGMPAPLGAT--ALAGGVNFAVYSGGATAAALCLFTPEDLKADRVTEEVPL 107
Db     25 ALLGAPAWAAINPNKLGAAYDATKANVTFKVYSSKATRIELYLYSTATGSAEKAKY---- 80

Qy    108 DPLMNRTGNVWHVFI-------EGELHNMLYGYRFDGTFAPH----------------- 142
Db     81 --VMTNSGGIWSVTIPTSTLSGQGLGGTLYYGYRAWGPNWPYNASWTKGSSLGFISDVDA 138

Qy    143 CGHYLDVSNVVVDPYA--------KAVISRGEY---GVPARGNNCWPQMAGMIPLPYSTF 191
Db    139 AGNRFNPNKLLSDPYALELSHDPTTATMTNGSIYASGATYRNIDSGSSAPKGIVLAGDT- 197

Qy    192 DWEGDLPLRYPQKDLVIYEMHLRGFTKHDSS-NVEHPGTFIGAVSKLDYLKELGVNCIEL 250
Db    198 QATGTKPTR-ALKDDVVYEAHVRGLTMNDTSITAAYRGTYKGAGLKAAALAALGVTAIEF 256

Qy    251 MPCHE----FNELEYSTSSSKMNFWGYSTINFFSPMTRY----TSGGIKNCGRDAINEFK 302
Db    257 LPVQETQNDTNDNDPSSTSGD-NYWGYMTLNYFAPDRRYAYDKTPGG-------PTREFK 308

Qy    303 TFVREAHKRGIEVILDVVFNHTAEGNENGP-------ILSFKGVDNTTYYMLAPKGEFYN 355
Db    309 EMVKAFHDNGIKVLVDVVYNHTGEGGAWSPTDKTTYNITSFRGLDNPTYYSLT--ADFQN 366

Qy    356 ---YSGCGNTFNCNHPVVRQFIVDCLRYWVTEMHVDGFRFDLASIMTRGSSLWDPVNVYG 412
Db    367 SWDNTGVGGNYNTRNTIAQNLIVDSLAYWRDKLGVDGYRFDLASVL--GNSC-----QHG 419

Qy    413 APIEGDMITTGTPLVTPPLIDMISNDPILG--GVKLIAEAWDAGG-LYQVGQFPHWNVWS 469
Db    420 C-FNFDKMDAGNAL--NRIVAELPPRPATGGSGVDLIAEPWAIGGNSYQVGGFP--SGWA 474

Qy    470 EWNGKYRDIVRQFIKGTDGFA---GGFAECLCGSPHLYQAGGRKPWHSINFVCAHDGFTL 526
Db    475 EWNGAYRDVVRQAQNKLGSVAITTGQMATRFAGSSDLYGDDGRKPWHSVNFITAHDGFTL 534

Qy    527 ADLVTYNKKYN---LPNGENNRDGENHNLSWNCGEEGEFARLSVKRLRKRQMRNFFVCLM 583
Db    535 KDLYSCNSKNNNQVWPYGPSD-GGEDNNNSWDQG--------GIAADQRKAARNGMALMM 585

Qy    584 VSQGVPMFYMGDEYGHTKGGNNNTYCHDSYVNYFRWDKKEQYSELHRFCCLMTKFRKECE 643
Db    586 LSAGVPMIVGGDEALRSMNCNNNPYNLDSSANWLNWSRTTDQNNFQSFSKAMIAFRKAHP 645

Qy    644 GLGLEDFPTA--------KRLQWHGHQPGKPDW---------SENSRFVAFSMKDERQGE 686
Db    646 ALRPANFYSSVDNNGNVMEQLRWF-----KPDGGVADATYFNDANNHAIAWRIDGSEFGD 700

Qy    687 ----IYVAFNTSHLPAVVELP-ERAGRRWEPVVDT 716
Db    701 TASAIYVAHNAWSAQVNFTLPWPGAGKSWYRVTDT 735


RESULT 11
US-08-476-519-2
; Sequence 2, Application US/08476519
; Patent No. 5750876
;  GENERAL INFORMATION:
;    APPLICANT:  Barry, Gerard F.
;    APPLICANT:  Kishore, Ganesh M.
;    APPLICANT:  Krohn, Bradley M.
;    TITLE OF INVENTION:  No. 5750876el Isoamylase Gene, Compositions
;    TITLE OF INVENTION:  Containing It and Methods of Using Isoamylases
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Grace L. Bonner, Monsanto Company, BB4F
;      STREET:  700 Chesterfield Parkway No. 5750876th
;      CITY:  St. Louis
;      STATE:  Missouri
;      COUNTRY:  USA
;      ZIP:  63198
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/476,519
;      FILING DATE:
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/281902
;      FILING DATE:  28-JUL-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bonner, Grace L.
;      REGISTRATION NUMBER:  32,963
;      REFERENCE/DOCKET NUMBER:  38-21(13577)A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314)537-7286
;      TELEFAX:  (314)537-6047
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  777 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-476-519-2
```

```
  Query Match             19.2%;  Score 810;  DB 1;  Length 777;
  Best Local Similarity   32.6%;  Pred. No. 4.4e-70;
  Matches  257;  Conservative   96;  Mismatches  271;  Indels  164;  Gaps   38;

Qy      42 EPVAEDRYALGGACRVL-------AGMPA-------PLGATALAGGVN--FAVYSGGATA 85
           :| |  |   |   | |       |  ||        |||   |   |  | |||  ||
Db       2 DPHAPQRQRSGQRLRALALAALACALSPAHAAIDAQQLGARYDAAQANLAFRVYSSRATR 61

Qy      86 AALCLF-TPED--------LKADRVTE----EVPLDPLMNRTGNVWHVFIEGELHNMLYG 132
             : |:  |           |  |  |:     :|   : |  |       | | ::   ||
Db      62 VEVFLYKNPTGSQEVARLALSKDPATQVWSLSLPTSTIKNTYG------ITGAVY---YG 112

Qy     133 YRFDGTFAPH---------CGHYLDVSN---------VVVDPYAK--------AVISRGE 166
           ||  |   |:          |   || |         :::||||:          |  : |
Db     113 YRAWGPNWPYDAAWTKGSATGFVSDVDNAGNRFNPNKLLLDPYAREISQDPNTATCADGT 172

Qy     167 -YGVPARGNN-----CWPQMAGMIPLPYSTFDWEGDLPLRYPQKDLVIYEMHLRGFTKHD 220
            |   |   |     |  :   :     |     | | |   || ||||:|:|| |::|
Db     173 IYATGAAHRNKDSGLCASKGIALAADATSV----GSKPTR-ALKDEVIYEVHVRGLTRND 227

Qy     221 SS-NVEHPGTFIGAVSKLDYLKELGVNCIELMPCHEF----NELEYSTSSSKMNFWGYST 275
           |      ||: ||  |   |  |||  :| :|  |     |:::   |::  |:||| |
Db     228 DSVPAAERGTYKGAARKAAALAALGVTAVEFLPVQETQNDQNDVD-PNSTAGDNYWGYMT 286

Qy     276 INFFSPMTRY----TSGGIKNCGRDAINEFKTFVREAHKRGIEVILDVVFNHTAEGNENG 331
           :|:|:|  ||    ::||          |:|  |:  |  ||:| :|||:||| ||   |
Db     287 LNYFAPDRRYAYDKSAGG-------PTREWKAMVKAFHDAGIKVYIDVVYNHTGEG---G 336

Qy     332 P-----------ILSFKGVDNTTYYMLAPKGEF-YNYSGCGNTFNCNHPVVRQFIVDCLR 379
           |           :|||:|:||  || |:   :: :: :| |  :|  || :  ||| |
Db     337 PWSGTDGLSVYNLLSFRGLDNPAYYSLSSDYKYPWDNTGVGGNYNTRHPIAQNLIVDSLA 396

Qy     380 YWVTEMHVDGFRFDLASIMTRGSSLWDPVNVYGAPIEGDMITTGTPLVTPPLIDMISNDP 439
           ||   : ||||||||||::  |:|       :     |: :          ::  :  |
Db     397 YWRDALGVDGFRFDLASVL--GNSCQHGCFNFDKNDSGNALNR--------IVAELPPRP 446

Qy     440 ILG--GVKLIAEAWDAGG-LYQVGQFPHWNVWSEWNGKYRDIVR--QFIKGTDGFA-GGF 493
             |  | |||| |  ||  |||| ||    |:|||| ||| :|  |    | :    |
Db     447 AAGGAGADLIAEPWAIGGNSYQVGGFPAG--WAEWNGLYRDALRKKQNKLGVETVTPGTL 504

Qy     494 AECLCGSPHLYQAGGRKPWHSINFVCAHDGFTLADLVTYNKKYN---LPNGENNRDGENH 550
           |    ||  ||   ||||||||||| |||||| ||  || | |    | | ::  ||:|
Db     505 ATRFAGSNDLYGDDGRKPWHSINFVVAHDGFTLNDLYAYNDKQNNQPWPYGPSD-GGEDH 563

Qy     551 NLSWNCGEEGEFARLSVKRLRKRQMRNFFVCLMVSQGVPMFYMGDEYGHTKGGNNNTYCH 610
           ||||| |         :   :::  |     ||:| ||||   |||   |: ||||||
Db     564 NLSWNQG--------GIVAEQRKAARTGLALLMLSAGVPMITGGDEALRTQFGNNNTYNL 615

Qy     611 DSYVNYFRWDKKEQYSELHRFCCLMTKFRKECEGLGLEDFPTA--------KRLQWHGHQ 662
           ||  |:  | :    ::   :   :  |||    |   :| :|         ::|:|
Db     616 DSAANWLYWSRSALEADHETYTKRLIAFRKAHPALRPANFYSASDTNGNVMEQLRWF--- 672

Qy     663 PGKPDWSE---------NSRFVAFSMKDERQGE----IYVAFNTSHLPAVVELP-ERAGR 708
             ||| ::          ::  :|: :     |:    ||||:|         :||    |:
Db     673 --KPDGAQADSAYFNGADNHALAWRIDGSEFGDSASAIYVAYNGWSGAVDFKLPWPGTGK 730

Qy     709 RWEPVVDT 716
           :|  | ||
Db     731 QWYRVTDT 738


RESULT  12
PCT-US95-09323-2
; Sequence 2, Application PC/TUS9509323
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  Novel Isoamylase Gene, Compositions
;    TITLE OF INVENTION:  Containing It and Methods of Using Isoamylases
;    NUMBER OF SEQUENCES:  11
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/09323
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/281902
;      FILING DATE:  28-JUL-1994
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  777 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
PCT-US95-09323-2


  Query Match             19.2%;  Score 810;  DB 5;  Length 777;
  Best Local Similarity   32.6%;  Pred. No. 4.4e-70;
  Matches  257;  Conservative   96;  Mismatches  271;  Indels  164;  Gaps   38;

Qy      42 EPVAEDRYALGGACRVL-------AGMPA-------PLGATALAGGVN--FAVYSGGATA 85
           :| |  |   |   | |       |  ||        |||   |   |  | |||  ||
Db       2 DPHAPQRQRSGQRLRALALAALACALSPAHAAIDAQQLGARYDAAQANLAFRVYSSRATR 61

Qy      86 AALCLF-TPED--------LKADRVTE----EVPLDPLMNRTGNVWHVFIEGELHNMLYG 132
             : |:  |           |  |  |:     :|   : |  |       | | ::   ||
Db      62 VEVFLYKNPTGSQEVARLALSKDPATQVWSLSLPTSTIKNTYG------ITGAVY---YG 112

Qy     133 YRFDGTFAPH---------CGHYLDVSN---------VVVDPYAK--------AVISRGE 166
           ||  |   |:          |   || |         :::||||:          |  : |
Db     113 YRAWGPNWPYDAAWTKGSATGFVSDVDNAGNRFNPNKLLLDPYAREISQDPNTATCADGT 172

Qy     167 -YGVPARGNN-----CWPQMAGMIPLPYSTFDWEGDLPLRYPQKDLVIYEMHLRGFTKHD 220
            |   |   |     |  :   :     |     | | |   || ||||:|:|| |::|
Db     173 IYATGAAHRNKDSGLCASKGIALAADATSV----GSKPTR-ALKDEVIYEVHVRGLTRND 227

Qy     221 SS-NVEHPGTFIGAVSKLDYLKELGVNCIELMPCHEF----NELEYSTSSSKMNFWGYST 275
           |      ||: ||  |   |  |||  :| :|  |     |:::   |::  |:||| |
Db     228 DSVPAAERGTYKGAARKAAALAALGVTAVEFLPVQETQNDQNDVD-PNSTAGDNYWGYMT 286

Qy     276 INFFSPMTRY----TSGGIKNCGRDAINEFKTFVREAHKRGIEVILDVVFNHTAEGNENG 331
           :|:|:|  ||    ::||          |:|  |:  |  ||:| :|||:||| ||   |
Db     287 LNYFAPDRRYAYDKSAGG-------PTREWKAMVKAFHDAGIKVYIDVVYNHTGEG---G 336

Qy     332 P-----------ILSFKGVDNTTYYMLAPKGEF-YNYSGCGNTFNCNHPVVRQFIVDCLR 379
           |           :|||:|:||  || |:   :: :: :| |  :|  || :  ||| |
Db     337 PWSGTDGLSVYNLLSFRGLDNPAYYSLSSDYKYPWDNTGVGGNYNTRHPIAQNLIVDSLA 396

Qy     380 YWVTEMHVDGFRFDLASIMTRGSSLWDPVNVYGAPIEGDMITTGTPLVTPPLIDMISNDP 439
           ||   : ||||||||||::  |:|       :     |: :          ::  :  |
Db     397 YWRDALGVDGFRFDLASVL--GNSCQHGCFNFDKNDSGNALNR--------IVAELPPRP 446

Qy     440 ILG--GVKLIAEAWDAGG-LYQVGQFPHWNVWSEWNGKYRDIVR--QFIKGTDGFA-GGF 493
             |  | |||| |  ||  |||| ||    |:|||| ||| :|  |    | :    |
Db     447 AAGGAGADLIAEPWAIGGNSYQVGGFPAG--WAEWNGLYRDALRKKQNKLGVETVTPGTL 504

Qy     494 AECLCGSPHLYQAGGRKPWHSINFVCAHDGFTLADLVTYNKKYN---LPNGENNRDGENH 550
           |    ||  ||   ||||||||||| |||||| ||  || | |    | | ::  ||:|
Db     505 ATRFAGSNDLYGDDGRKPWHSINFVVAHDGFTLNDLYAYNDKQNNQPWPYGPSD-GGEDH 563

Qy     551 NLSWNCGEEGEFARLSVKRLRKRQMRNFFVCLMVSQGVPMFYMGDEYGHTKGGNNNTYCH 610
           ||||| |         :   :::  |     ||:| ||||   |||   |: ||||||
Db     564 NLSWNQG--------GIVAEQRKAARTGLALLMLSAGVPMITGGDEALRTQFGNNNTYNL 615

Qy     611 DSYVNYFRWDKKEQYSELHRFCCLMTKFRKECEGLGLEDFPTA--------KRLQWHGHQ 662
           ||  |:  | :    ::   :   :  |||    |   :| :|         ::|:|
Db     616 DSAANWLYWSRSALEADHETYTKRLIAFRKAHPALRPANFYSASDTNGNVMEQLRWF--- 672

Qy     663 PGKPDWSE---------NSRFVAFSMKDERQGE----IYVAFNTSHLPAVVELP-ERAGR 708
             ||| ::          ::  :|: :     |:    ||||:|         :||    |:
```

```
Db    673 --KPDGAQADSAYFNGADNHALAWRIDGSEFGDSASAIYVAYNGWSGAVDFKLPWPGTGK 730

Qy    709 RWEPVVDT 716
          :|  | ||
Db    731 QWYRVTDT 738


RESULT  13
US-08-476-519-11
; Sequence 11, Application US/08476519
; Patent No. 5750876
;  GENERAL INFORMATION:
;    APPLICANT:  Barry, Gerard F.
;    APPLICANT:  Kishore, Ganesh M.
;    APPLICANT:  Krohn, Bradley M.
;    TITLE OF INVENTION:  No. 5750876el Isoamylase Gene, Compositions
;    TITLE OF INVENTION:  Containing It and Methods of Using Isoamylases
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Grace L. Bonner, Monsanto Company, BB4F
;      STREET:  700 Chesterfield Parkway No. 5750876th
;      CITY:  St. Louis
;      STATE:  Missouri
;      COUNTRY:  USA
;      ZIP:  63198
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/476,519
;      FILING DATE:
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/281902
;      FILING DATE:  28-JUL-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bonner, Grace L.
;      REGISTRATION NUMBER:  32,963
;      REFERENCE/DOCKET NUMBER:  38-21(13577)A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314)537-7286
;      TELEFAX:  (314)537-6047
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  746 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-476-519-11


  Query Match           19.2%;  Score 807;  DB 1;  Length 746;
  Best Local Similarity   33.0%;  Pred. No. 8.1e-70;
  Matches  248;  Conservative   95;  Mismatches  258;  Indels  150;  Gaps   36;

Qy     65 LGATALAGGVN--FAVYSGGATAAALCLF-TPED--------LKADRVTE----EVPLDP 109
          |||   |   | | |||  ||   : |:  |           |  |  |:     :|
Db      8 LGARYDAAQANLAFRVYSSRATRVEVFLYKNPTGSQEVARLALSKDPATQVWSLSLPTST 67

Qy    110 LMNRTGNVWHVFIEGELHNMLYGYRFDGTFAPH---------CGHYLDVSN--------- 151
          : |  |      | | ::   ||||  |   |:           |   || |
Db     68 IKNTYG------ITGAVY---YGYRAWGPNWPYDAAWTKGSATGFVSDVDNAGNRFNPNK 118

Qy    152 VVVDPYAK--------AVISRGE-YGVPARGNN-----CWPQMAGMIPLPYSTFDWEGDL 197
          :::||||:        |  : |  |   |   |     :     :     |     |
Db    119 LLLDPYAREISQDPNTATCADGTIYATGAAHRNKDSGLCASKGIALAADATSV----GSK 174

Qy    198 PLRYPQKDLVIYEMHLRGFTKHDSS-NVEHPGTFIGAVSKLDYLKELGVNCIELMPCHEF 256
          | |   || |||||:|:|| |::| |      ||: ||  |   |  |||  :| :|  |
Db    175 PTR-ALKDEVIYEVHVRGLTRNDDSVPAAERGTYKGAARKAAALAALGVTAVEFLPVQET 233

Qy    257 ----NELEYSTSSSKMNFWGYSTINFFSPMTRY----TSGGIKNCGRDAINEFKTFVREA 308
              |:::   |::  |:||| |:|:|:|  ||    ::||          |:|  |:
Db    234 QNDQNDVD-PNSTAGDNYWGYMTLNYFAPDRRYAYDKSAGG-------PTREWKAMVKAF 285

Qy    309 HKRGIEVILDVVFNHTAEGNENGP----------ILSFKGVDNTTYYMLAPKGEF-YNY 356
          |  ||:| :|||:||| ||   ||           :|||:|:|| || |:   :: ::
Db    286 HDAGIKVYIDVVYNHTGEG---GPWSGTDGLSVYNLLSFRGLDNPAYYSLSSDYKYPWDN 342

Qy    357 SGCGNTFNCNHPVVRQFIVDCLRYWVTEMHVDGFRFDLASIMTRGSSLWDPVNVYGAPIE 416
          :| |  :|  ||: :  ||| | ||   : ||||||||||::  |:|       :
Db    343 TGVGGNYNTRHPIAQNLIVDSLAYWRDALGVDGFRFDLASVL--GNSCQHGCFNFDKNDS 400

Qy    417 GDMITTGTPLVTPPLIDMISNDPILG--GVKLIAEAWDAGG-LYQVGQFPHWNVWSEWNG 473
          |: :           ::  :   |  |  | |||| |  ||  |||| ||    |:||||
Db    401 GNALNR--------IVAELPPRPAAGGAGADLIAEPWAIGGNSYQVGGFPAG--WAEWNG 450

Qy    474 KYRDIVR--QFIKGTDGFA-GGFAECLCGSPHLYQAGGRKPWHSINFVCAHDGFTLADLV 530
          ||| :|  |   | :    |  |    ||  ||   |||||||||||| ||||||| ||
Db    451 LYRDALRKKQNKLGVETVTPGTLATRFAGSNDLYGDDGRKPWHSINFVVAHDGFTLNDLY 510

Qy    531 TYNKKYN---LPNGENNRDGENHNLSWNCGEEGEFARLSVKRLRKRQMRNFFVCLMVSQG 587
          || | |    | | ::  ||:|||||| |          :  :::  |    ||:| |
Db    511 AYNDKQNNQPWPYGPSD-GGEDHNLSWNQG--------GIVAEQRKAARTGLALLMLSAG 561

Qy    588 VPMFYMGDEYGHTKGGNNNTYCHDSYVNYFRWDKKEQYSELHRFCCLMTKFRKECEGLGL 647
          |||   |||   |: ||||||  ||  |:  | :    ::   :   :  |||    |
Db    562 VPMITGGDEALRTQFGNNNTYNLDSAANWLYWSRSALEADHETYTKRLIAFRKAHPALRP 621

Qy    648 EDFPTA--------KRLQWHGHQPGKPDWSE---------NSRFVAFSMKDERQGE---- 686
          :| :|        ::|:|      ||| ::          ::  :|: :     |:
Db    622 ANFYSASDTNGNVMEQLRWF-----KPDGAQADSAYFNGADNHALAWRIDGSEFGDSASA 676

Qy    687 IYVAFNTSHLPAVVELP-ERAGRRWEPVVDT 716
          ||||:|        :||    |::|  | ||
Db    677 IYVAYNGWSGAVDFKLPWPGTGKQWYRVTDT 707


RESULT  14
PCT-US95-09323-11
; Sequence 11, Application PC/TUS9509323
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  Novel Isoamylase Gene, Compositions
;    TITLE OF INVENTION:  Containing It and Methods of Using Isoamylases
;    NUMBER OF SEQUENCES:  11
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/09323
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/281902
;      FILING DATE:  28-JUL-1994
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  746 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
PCT-US95-09323-11


  Query Match           19.2%;  Score 807;  DB 5;  Length 746;
  Best Local Similarity   33.0%;  Pred. No. 8.1e-70;
  Matches  248;  Conservative   95;  Mismatches  258;  Indels  150;  Gaps   36;

Qy     65 LGATALAGGVN--FAVYSGGATAAALCLF-TPED--------LKADRVTE----EVPLDP 109
          |||   |   | | |||  ||   : |:  |           |  |  |:     :|
Db      8 LGARYDAAQANLAFRVYSSRATRVEVFLYKNPTGSQEVARLALSKDPATQVWSLSLPTST 67
```

```
Qy    110 LMNRTGNVWHVFIEGELHNMLYGYRFDGTFAPH---------CGHYLDVSN--------- 151
Db     68 IKNTYG------ITGAVY---YGYRAWGPNWPYDAAWTKGSATGFVSDVDNAGNRFNPNK 118

Qy    152 VVVDPYAK--------AVISRGE-YGVPARGNN-----CWPQMAGMIPLPYSTFDWEGDL 197
Db    119 LLLDPYAREISQDPNTATCADGTIYATGAAHRNKDSGLCASKGIALAADATSV----GSK 174

Qy    198 PLRYPQKDLVIYEMHLRGFTKHDSS-NVEHPGTFIGAVSKLDYLKELGVNCIELMPCHEF 256
Db    175 PTR-ALKDEVIYEVHVRGLTRNDDSVPAAERGTYKGAARKAAALAALGVTAVEFLPVQET 233

Qy    257 ----NELEYSTSSSKMNFWGYSTINFFSPMTRY----TSGGIKNCGRDAINEFKTFVREA 308
Db    234 QNDQNDVD-PNSTAGDNYWGYMTLNYFAPDRRYAYDKSAGG-------PTREWKAMVKAF 285

Qy    309 HKRGIEVILDVVFNHTAEGNENGP-----------ILSFKGVDNTTYYMLAPKGEF-YNY 356
Db    286 HDAGIKVYIDVVYNHTGEG---GPWSGTDGLSVYNLLSFRGLDNPAYYSLSSDYKYPWDN 342

Qy    357 SGCGNTFNCNHPVVRQFIVDCLRYWVTEMHVDGFRFDLASIMTRGSSLWDPVNVYGAPIE 416
Db    343 TGVGGNYNTRHPIAQNLIVDSLAYWRDALGVDGFRFDLASVL--GNSCQHGCFNFDKNDS 400

Qy    417 GDMITTGTPLVTPPLIDMISNDPILG--GVKLIAEAWDAGG-LYQVGQFPHWNVWSEWNG 473
Db    401 GNALNR--------IVAELPPRPAAGGAGADLIAEPWAIGGNSYQVGGFPAG--WAEWNG 450

Qy    474 KYRDIVR--QFIKGTDGFA-GGFAECLCGSPHLYQAGGRKPWHSINFVCAHDGFTLADLV 530
Db    451 LYRDALRKKQNKLGVETVTPGTLATRFAGSNDLYGDDGRKPWHSINFVVAHDGFTLNDLY 510

Qy    531 TYNKKYN---LPNGENNRDGENHNLSWNCGEEGEFARLSVKRLRKRQMRNFFVCLMVSQG 587
Db    511 AYNDKQNNQPWPYGPSD-GGEDHNLSWNQG--------GIVAEQRKAARTGLALLMLSAG 561

Qy    588 VPMFYMGDEYGHTKGGNNNTYCHDSYVNYFRWDKKEQYSELHRFCCLMTKFRKECEGLGL 647
Db    562 VPMITGGDEALRTQFGNNNTYNLDSAANWLYWSRSALEADHETYTKRLIAFRKAHPALRP 621

Qy    648 EDFPTA--------KRLQWHGHQPGKPDWSE---------NSRFVAFSMKDERQGE---- 686
Db    622 ANFYSASDTNGNVMEQLRWF-----KPDGAQADSAYFNGADNHALAWRIDGSEFGDSASA 676

Qy    687 IYVAFNTSHLPAVVELP-ERAGRRWEPVVDT 716
Db    677 IYVAYNGWSGAVDFKLPWPGTGKQWYRVTDT 707


RESULT 15
US-09-346-237-7
; Sequence 7, Application US/09346237A
; Patent No. 6265197
; GENERAL INFORMATION:
;  APPLICANT: Bisgard-Frantzen, Henrik
;  APPLICANT: Svendsen, Allan
;  TITLE OF INVENTION: Starch Debranching Enzymes
;  FILE REFERENCE: 5629.200-US
;  CURRENT APPLICATION NUMBER: US/09/346,237A
;  CURRENT FILING DATE: 1999-07-01
;  EARLIER APPLICATION NUMBER: PA 1998 00868
;  EARLIER FILING DATE: 1998-07-02
;  EARLIER APPLICATION NUMBER: 60/094,353
;  EARLIER FILING DATE: 1998-07-28
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 7
;  LENGTH: 776
;  TYPE: PRT
;  ORGANISM: Pseudomonas species SMP1
;  FEATURE:
;  NAME/KEY: PEPTIDE
;  LOCATION: (1)..(776)
;  OTHER INFORMATION: Isoamylase
US-09-346-237-7


  Query Match             19.1%;  Score 803.5;  DB 4;  Length 776;
  Best Local Similarity   32.0%;  Pred. No. 1.9e-69;
  Matches  242;  Conservative   99;  Mismatches  289;  Indels  127;  Gaps   31;

Qy     55 CRVLAGMPA----------PLGAT--ALAGGVNFAVYSGGATAAALCLFTPEDLKADRVT 102
Db     13 CAVLAGVPAMPAHAAINSMSLGASYDAQQANITFRVYSSQATRIVLYLYS-----AGYGV 67

Qy    103 EEVPLDPLMNRTGNVWHVFI-------EGELHNMLYGYRFDGTFAPHC----------- 143
Db     68 QESATYTLSPAGSGVWAVTVPVSSIKAAGITGAVYYGYRAWGPNWPYASNWGKGSQAGFV 127

Qy    144 ------GHYLDVSNVVVDPYAKAVISRGEYGVPARGNNCW--------------PQMAG 182
Db    128 SDVDANGDRFNPNKLLLDPYAQEV-SQDPLNPSNQNGNVFASGASYRTTDSGIYAPKGVV 186

Qy    183 MIPLPYSTFDWEGDLPLRYPQKDLVIYEMHLRGFTKHDSS-NVEHPGTFIGAVSKLDYLK 241
Db    187 LVPSTQST----GTKPTR-AQKDDVIYEVHVRGFTEQDTSIPAQYRGTYYGAGLKASYLA 241

Qy    242 ELGVNCIELMPCHE----FNELEYSTSSSKMNFWGYSTINFFSPMTRY----TSGGIKNC 293
Db    242 SLGVTAVEFLPVQETQNDANDV-VPNSDANQNYWGYMTENYFSPDRRYAYNKAAGG---- 296

Qy    294 GRDAINEFKTFVREAHKRGIEVILDVVFNHTAEGN-------ENGPILSFKGVDNTTYYM 346
Db    297 ---PTAEFQAMVQAFHNAGIKVYMDVVYNHTAEGGTWTSSDPTTATIYSWRGLDNTTYYE 353

Qy    347 LAPKGE-FYNYSGCGNTFNCNHPVVRQFIVDCLRYWVTEMHVDGFRFDLASIMTRGSSLW 405
Db    354 LTSGNQYFYDNTGIGANFNTYNTVAQNLIVDSLAYWANTMGVDGFRFDLASVL--GNSCL 411

Qy    406 DPVNVYGAP--IEG--DMITTGTPLVTPPLIDMISNDPILG--GVKLIAEAWDAGG-LYQ 458
Db    412 NGAYTASAPNCPNGGYNFDAADSNVAINRILREFTVRPAAGGSGLDLFAEPWAIGGNSYQ 471

Qy    459 VGQFPHWNVWSEWNGKYRDIVRQF---IKGTDGFAGGFAECLCGSPHLYQAGGRKPWHSI 515
Db    472 LGGFPQG--WSEWNGLFRDSLRQAQNELGSMTIYVTQDANDFSGSSNLFQSSGRSPWNSI 529

Qy    516 NFVCAHDGFTLADLVTYNKKYN---LPNGENNRDGENHNLSWNCGEEGEFARLSVKRLRK 572
Db    530 NFIDVHDGMTLKDVYSCNGANNSQAWPYGPSD-GGTSTNYSW---DQGMSAGTGAAVDQR 585

Qy    573 RQMRNFFVCLMVSQGVPMFYMGDEYGHTKGGNNNTYCHDSYVNYFRWDKKEQYSELHRFC 632
Db    586 RAARTGMAFEMLSAGTPLMQGGDEYLRTLQCNNNAYNLDSSANWLTYSWTTDQSNFYTFA 645

Qy    633 CLMTKFRKECEGLGLEDFPTAKRLQWHGHQPGKPD---WSENSRF-VAFSMKDERQGE-- 686
Db    646 QRLIAFRKAHPALRPSSWYSGSQLTWYQPSGAVADSNYWNNTSNYAIAYAINGPSLGDSN 705

Qy    687 -IYVAFN------TSHLPAVVELPERAGRRWEPVVDT 716
Db    706 SIYVAYNGWSSSVTFTLPA-----PPSGTQWYRVTDT 737


Search completed: July 31, 2002, 12:06:59
Job time: 399 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: July 31, 2002, 12:05:30 ; Search time 44.53 Seconds
(without alignments)
1648.602 Million cell updates/sec

Title: US-09-674-817A-2
Perfect score: 4212
Sequence: 1 SGPAPRLRRWRPNATAGKGV.....SNLYPMLSYSSVILVLRPDV 764

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3568.5	84.7	789	2	T01321	probable isoamylas
2	3527.5	83.7	818	2	T02231	probable isoamylas
3	2895.5	68.7	783	2	B84823	probable isoamylas
4	1638.5	38.9	707	2	S77094	glycogen operon pr
5	1538.5	36.5	702	2	B85091	isoamylase-like pr
6	1388	33.0	746	2	S74546	glycogen operon pr
7	1305.5	31.0	664	2	B86539	glycogen hydrolase
8	1305.5	31.0	664	2	B72084	glycosyl hydrolase
9	1300	30.9	666	2	E71565	probable glycogen
10	1268	30.1	720	2	B75540	glycogen operon pr
11	1264	30.0	666	2	G81717	glycosyl hydrolase
12	1225	29.1	721	2	A70764	probable glgX prot
13	1211	28.8	718	2	S73088	glycogen operon pr
14	1149.5	27.3	716	2	C83375	probable glycosyl
15	1123.5	26.7	698	2	C96001	probable glycosyl
16	1101.5	26.2	603	2	AC0674	probable glycogen
17	1014	24.1	659	2	A64119	glyX protein (EC 3
18	1005.5	23.9	657	2	JC7767	isoamylase (EC 3.2
19	1002	23.8	657	2	D91163	glycogen operon pr
20	1001	23.8	657	1	BVECGX	glycogen operon pr
21	997	23.7	657	2	E86009	hypothetical prote
22	980.5	23.3	654	2	AG0995	glycogen operon pr
23	967.5	23.0	662	2	AG0479	probable alpha-amy
24	927.5	22.0	656	2	D82388	glycogen operon pr
25	890	21.1	882	2	F86164	hypothetical prote
26	815.5	19.4	651	2	B98229	glycogen debranchi
27	815.5	19.4	651	2	AD3057	glycogen debranchi
28	803.5	19.1	776	2	A37035	isoamylase (EC 3.2
29	798.5	19.0	776	2	S13470	isoamylase (EC 3.2
30	527.5	12.5	718	2	G69585	pullulanase amyX -
31	523	12.4	843	2	H72204	pullulanase - Ther
32	519.5	12.3	720	2	G97229	pullulanase [impor
33	505.5	12.0	717	2	H84057	pullulanase amyX [
34	420	10.0	759	2	A98000	alpha-dextrin endo
35	418	9.9	759	2	D95129	pullulanase, proba
36	408	9.7	1072	2	A84112	alkaline amylopull
37	395.5	9.4	1256	2	G97902	alpha-amylase (EC
38	394.5	9.4	1280	2	E95031	alkaline amylopull
39	370.5	8.8	561	2	S73087	alpha-amylase (EC
40	354	8.4	691	2	D90592	hypothetical prote
41	344	8.2	964	2	S51324	pullulanase - spin
42	335	8.0	558	2	JC5135	alpha-amylase (EC
43	327.5	7.8	910	2	D75524	alpha-dextran endo
44	327	7.8	986	2	T03760	alpha-dextrin endo
45	325.5	7.7	904	2	T04377	probable alpha-dex

ALIGNMENTS

RESULT 1
T01321
probable isoamylase (EC 3.2.1.68) sul - maize
C;Species: Zea mays (maize)
C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 29-Oct-1999
C;Accession: T01321
R;Park, S.Y.; Kim, Y.H.
Plant Physiol. 115, 1731, 1997
A;Title: Cloning and nucleotide sequencing of a cDNA encoding flavanone 3-hydroxyla
A;Reference number: Z14292
A;Accession: T01321
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-789 <PAR>
A;Cross-references: EMBL:AF030882; NID:g2623247; PIDN:AAB97167.1; PID:g2623248
C;Genetics:
A;Gene: sugary1
A;Introns: 121/3; 183/3; 214/3; 265/3; 291/2; 338/3; 374/3; 403/2; 427/2; 469/3; 49
C;Function:
A;Description: starch debranching enzyme
C;Superfamily: glyX protein
C;Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 84.7%; Score 3568.5; DB 2; Length 789;
Best Local Similarity 83.6%; Pred. No. 1.3e-266;
Matches 654; Conservative 44; Mismatches 65; Indels 19; Gaps 6;

```
Qy       1 SGPAPRLR----RW------RPNATAGKGVGEV------CAAVVEAATKVEDEGEEDEPV 44
           | | | |     ||      |||  || | | :       | ||    ||: ::|| |
Db       9 SSPRPLLAVPAGRWRAGVRGRPN-VAGLGRGRLSLHAAAARPVAEAVQAEEDDDDDDEEV 67

Qy      45 AEDRYALGGACRVLAGMPAPLGATALAGGVNFAVYSGGATAAALCLFTPEDLKADRVTEE 104
           ||:|:||||||||||||||||||| |||||||| ||:||:|||| | ||||||||||
Db      68 AEERFALGGACRVLAGMPAPLGATALRGGVNFAVYSSGASAASLCLFAPGDLKADRVTEE 127

Qy     105 VPLDPLMNRTGNVWHVFIEG-ELHNMLYGYRFDGTFAPHCGHYLDVSNVVVDPYAKAVIS 163
           ||||||:|||||||||||| | :|| |||||||| |||  | | |||||||||||||:|
Db     128 VPLDPLLNRTGNVWHVFIHGDQLHGMLYGYRFDGVFAPERGQYYDVSNVVVDPYAKAVVS 187

Qy     164 RGEYGVPARGNNCWPQMAGMIPLPYSTFDWEGDLPLRYPQKDLVIYEMHLRGFTKHDSSN 223
           |||||||| | :|||||||||||||: |||:|||| | ||||||||||||||||:||
Db     188 RGEYGVPAPGGSCWPQMAGMIPLPYNKFDWQGDLPLGYHQKDLVIYEMHLRGFTKHNSSK 247

Qy     224 VEHPGTFIGAVSKLDYLKELGVNCIELMPCHEFNELEYSTSSSKMNFWGYSTINFFSPMT 283
           :||||:|||||||:|||||||||||||||||||| :|||||||||||||||||
Db     248 TKHPGTYIGAVSKLDHLKELGVNCIELMPCHEFNELEYFSSSSKMNFWGYSTINFFSPMA 307

Qy     284 RYTSGGIKNCGRDAINEFKTFVREAHKRGIEVILDVVFNHTAEGNENGPILSFKGVDNTT 343
           ||:| ||:: |  |||||| ||||||||||||:|||||||||||| ||||||:|:||:|
Db     308 RYSSSGIRDSGCGAINEFKAFVREAHKRGIEVIMDVVFNHTAEGNEKGPILSFRGIDNST 367
```

```
Qy    344 YYMLAPKGEFYNYSGCGNTFNCNHPVVRQFIVDCLRYWVTEMHVDGFRFDLASIMTRGSS 403
          ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||:||| |
Db    368 YYMLAPKGEFYNYSGCGNTFNCNHPVVREFIVDCLRYWVTEMHVDGFRFDLASILTRGCS 427

Qy    404 LWDPVNVYGAPIEGDMITTGTPLVTPPLIDMISNDPILGGVKLIAEAWDAGGLYQVGQFP 463
          |||||||||:|:||||||||||||| ||||||||||||| ||||||||||||||||||||
Db    428 LWDPVNVYGSPMEGDMITTGTPLVAPPLIDMISNDPILGNVKLIAEAWDAGGLYQVGQFP 487

Qy    464 HWNVWSEWNGKYRDIVRQFIKGTDGFAGGFAECLCGSPHLYQAGGRKPWHSINFVCAHDG 523
          |||||||||||||| |||||||||||| |||||||| ||||||||||||||||||||||||
Db    488 HWNVWSEWNGKYRDTVRQFIKGTDGFAGAFAECLCGSPQLYQAGGRKPWHSINFVCAHDG 547

Qy    524 FTLADLVTYNKKYNLPNGENNRDGENHNLSWNCGEEGEFARLSVKRLRKRQMRNFFVCLM 583
          |||||||||| |||| |||:|||||||||||||||||||||| |||:|||||||||||||||
Db    548 FTLADLVTYNSKYNLSNGEDNRDGENHNLSWNCGEEGEFASLSVRRLRKRQMRNFFVCLM 607

Qy    584 VSQGVPMFYMGDEYGHTKGGNNNTYCHDSYVNYFRWDKK-EQYSELHRFCCLMTKFRKEC 642
          |||||||||||||||||||||||||||||| |||||||||| || |:|:||| |||||||||
Db    608 VSQGVPMFYMGDEYGHTKGGNNNTYCHDHYVNYFRWDKKEEQSSDLYRFCRLMTKFRKEC 667

Qy    643 EGLGLEDFPTAKRLQWHGHQPGKPDWSENSRFVAFSMKDERQGEIYVAFNTSHLPAVVEL 702
          | |||||||::||:|||||||||||||| ||||||:|||| :|||||||||||| || |
Db    668 ESLGLEDFPTSERLKWHGHQPGKPDWSEASRFVAFTMKDETKGEIYVAFNTSHLPVVVGL 727

Qy    703 PERAGRRWEPVVDTGKPAPYDFLTDDLPDRALTIHQFSHFLYSNLYPMLSYSSVILVLRP 762
          |||:| |||||||||| |||||||| |||||:|::|||||| ||||||||||:||||||
Db    728 PERSGFRWEPVVDTGKEAPYDFLTDGLPDRAVTVYQFSHFLNSNLYPMLSYSSIILVLRP 787

Qy    763 DV 764
          ||
Db    788 DV 789
```

RESULT 2
T02231
probable isoamylase (EC 3.2.1.68) su1 - maize (fragment)
C;Species: Zea mays (maize)
C;Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 29-Oct-1999
C;Accession: T02231
R;James, M.G.; Robertson, D.S.; Myers, A.M.
Plant Cell 7, 417-429, 1995
A;Title: Characterization of the maize gene sugary1, a determinant of starch compositior
A;Reference number: Z14630; MUID:95290995
A;Accession: T02231
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-818 <JAM>
A;Cross-references: EMBL:U18908; NID:g758725; PIDN:AAA91298.1; PID:g758726
A;Experimental source: endosperm
C;Genetics:
A;Gene: sugary1
A;Map position: 4s
C;Function:
A;Description: starch debranching enzyme
C;Superfamily: glyX protein
C;Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 83.7%; Score 3527.5; DB 2; Length 818;
Best Local Similarity 83.0%; Pred. No. 2e-263;
Matches 649; Conservative 44; Mismatches 70; Indels 19; Gaps 6;

```
Qy      1 SGPAPRLR----RW------RPNATAGKGVGEV------CAAVVEAATKVEDEGEEDEPV 44
          | | | |     ||      |||  || | | :         | ||    ||: ::|| |
Db     38 SSPRPLLAVPAGRWRAGVRGRPN-VAGLGRGRLSLHAAAARPVAEAVQAEEDDDDDDEEV 96

Qy     45 AEDRYALGGACRVLAGMPAPLGATALAGGVNFAVYSGGATAAALCLFTPEDLKADRVTEE 104
          ||:|:|||||||||||||||||||| |||||||| ||:||:| || | ||||||||||
Db     97 AEERFALGGACRVLAGMPAPLGATALRGGVNFAVYSSGASAASLSLFAPGDLKADRVTEE 156

Qy    105 VPLDPLMNRTGNVWHVFIEG-ELHNMLYGYRFDGTFAPHCGHYLDVSNVVVDPYAKAVIS 163
          ||||||:|||||||||||| | ||| || |||||| |||  | | |||||||||||||:|
Db    157 VPLDPLLNRTGNVWHVFIHGDELHGMLCGYRFDGVFAPERGQYYDVSNVVVDPYAKAVVS 216
```

```
Qy    164 RGEYGVPARGNNCWPQMAGMIPLPYSTFDWEGDLPLRYPQKDLVIYEMHLRGFTKHDSSN 223
          |||||||| | :|||||||||||||: |||:||||| | ||||||||||||||||:||
Db    217 RGEYGVPAPGGSCWPQMAGMIPLPYNKFDWQGDLPLGYHQKDLVIYEMHLRGFTKHNSSK 276

Qy    224 VEHPGTFIGAVSKLDYLKELGVNCIELMPCHEFNELEYSTSSSKMNFWGYSTINFFSPMT 283
          :||||:|||||||:|||||||||||||||||||||| :|||||||||||||||||||
Db    277 TKHPGTYIGAVSKLDHLKELGVNCIELMPCHEFNELEYFSSSSKMNFWGYSTINFFSPMA 336

Qy    284 RYTSGGIKNCGRDAINEFKTFVREAHKRGIEVILDVVFNHTAEGNENGPILSFKGVDNTT 343
          ||:| ||:: |  |||||| ||||||||||||:|||||||||||| |||||:|:||:|
Db    337 RYSSSGIRDSGCGAINEFKAFVREAHKRGIEVIMDVVFNHTAEGNEKGPILSFRGIDNST 396

Qy    344 YYMLAPKGEFYNYSGCGNTFNCNHPVVRQFIVDCLRYWVTEMHVDGFRFDLASIMTRGSS 403
          ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||:||| |
Db    397 YYMLAPKGEFYNYSGCGNTFNCNHPVVREFIVDCLRYWVTEMHVDGFRFDLASILTRGCS 456

Qy    404 LWDPVNVYGAPIEGDMITTGTPLVTPPLIDMISNDPILGGVKLIAEAWDAGGLYQVGQFP 463
          |||||||||:|:||||||||||||| ||||||||||||| |||||||||||||||| ||||
Db    457 LWDPVNVYGSPMEGDMITTGTPLVAPPLIDMISNDPILGNVKLIAEAWDAGGLYQEGQFP 516

Qy    464 HWNVWSEWNGKYRDIVRQFIKGTDGFAGGFAECLCGSPHLYQAGGRKPWHSINFVCAHDG 523
          |||||||||||||| |||||||||||| |||||||| ||||||||||||||||||||||||
Db    517 HWNVWSEWNGKYRDTVRQFIKGTDGFAGAFAECLCGSPQLYQAGGRKPWHSIGFVCAHDG 576

Qy    524 FTLADLVTYNKKYNLPNGENNRDGENHNLSWNCGEEGEFARLSVKRLRKRQMRNFFVCLM 583
          |||||||||| |||| |||: |||||||||||||||||||| |||:|||||||||||||||
Db    577 FTLADLVTYNSKYNLSNGEDFRDGENHNLSWNCGEEGEFASLSVRRLRKRQMRNFFVCLM 636

Qy    584 VSQGVPMFYMGDEYGHTKGGNNNTYCHDSYVNYFRWDKK-EQYSELHRFCCLMTKFRKEC 642
          |||||||||||||||||||||||||||||| |||||||||| || |:|:||| |||:|||||
Db    637 VSQGVPMFYMGDEYGHTKGGNNNTYCHDHYVNYFRWDKKEEQSSDLYRFCRLMTEFRKEC 696

Qy    643 EGLGLEDFPTAKRLQWHGHQPGKPDWSENSRFVAFSMKDERQGEIYVAFNTSHLPAVVEL 702
          | |||||||::||:|||||||||||||| ||||||:|||| :|||||||||||| || |
Db    697 ESLGLEDFPTSERLKWHGHQPGKPDWSEASRFVAFTMKDETKGEIYVAFNTSHLPVVVGL 756

Qy    703 PERAGRRWEPVVDTGKPAPYDFLTDDLPDRALTIHQFSHFLYSNLYPMLSYSSVILVLRP 762
          |||:| |||||||||| |||||||| |||||:|::|||||| ||||||||||:||||||
Db    757 PERSGFRWEPVVDTGKEAPYDFLTDGLPDRAVTVYQFSHFLNSNLYPMLSYSSIILVLRP 816

Qy    763 DV 764
          ||
Db    817 DV 818
```

RESULT 3
B84823
probable isoamylase [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C;Accession: B84823
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Vente
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487
A;Accession: B84823
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-783 <STO>
A;Cross-references: GB:AE002093; NID:g2088646; PIDN:AAB95278.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g39930
A;Map position: 2
C;Superfamily: glyX protein

Query Match 68.7%; Score 2895.5; DB 2; Length 783;
Best Local Similarity 69.3%; Pred. No. 9.2e-215;
Matches 516; Conservative 84; Mismatches 128; Indels 17; Gaps 5;

```
Qy      28 VEAATKVEDEGEE----DEPVAEDRYALGGACRVLAGMPAPLGATALAGGVNFAVYSGGA 83
Db      42 ISAKDRRSNEAENIAVVEKPLKSDRFFISD------GLPSPFGPTVRDDGVNFSVYSTNS 95

Qy      84 TAAALCLFTPEDLKADRVTEEVPLDPLMNRTGNVWHVFIEGELHNMLYGYRFDGTFAPHC 143
Db      96 VSATICLISLSDLRQNKVTEEIQLDPSRNRTGHVWHVFLRGDFKDMLYGYRFDGKFSPEE 155

Qy     144 GHYLDVSNVVVDPYAKAVISRGEYGVPARGNNCWPQMAGMIPLPYSTFDWEGDLPLRYPQ 203
Db     156 GHYYDSSNILLDPYAKAIISRDEFGVLGPDDNCWPQMACMVPTREEEFDWEGDMHLKLPQ 215

Qy     204 KDLVIYEMHLRGFTKHDSSNVEHPGTFIGAVSKLDYLKELGVNCIELMPCHEFNELEYST 263
Db     216 KDLVIYEMHVRGFTRHESSKIEFPGTYQGVAEKLDHLKELGINCIELMPCHEFNELEYYS 275

Qy     264 SSS-----KMNFWGYSTINFFSPMTRYTSGGIKNCGRDAINEFKTFVREAHKRGIEVILD 318
Db     276 YNTILGDHRVNFWGYSTIGFFSPMIRYASASSNNFAGRAINEFKILVKEAHKRGIEVIMD 335

Qy     319 VVFNHTAEGNENGPILSFKGVDNTTYYMLAPKGEFYNYSGCGNTFNCNHPVVRQFIVDCL 378
Db     336 VVLNHTAEGNEKGPIFSFRGVDNSVYYMLAPKGEFYNYSGCGNTFNCNHPVVRQFILDCL 395

Qy     379 RYWVTEMHVDGFRFDLASIMTRGSSLWDPVNVYGAPIEGDMITTGTPLVTPPLIDMISND 438
Db     396 RYWVTEMHVDGFRFDLGSIMSRSSSLWDAANVYGADVEGDLLTTGTPISCPPVIDMISND 455

Qy     439 PILGGVKLIAEAWDAGGLYQVGQFPHWNVWSEWNGKYRDIVRQFIKGTDGFAGGFAECLC 498
Db     456 PILRGVKLIAEAWDAGGLYQVGMFPHWGIWSEWNGKFRDVVRQFIKGTDGFSGAFAECLC 515

Qy     499 GSPHLYQAGGRKPWHSINFVCAHDGFTLADLVTYNKKYNLPNGENNRDGENHNLSWNCGE 558
Db     516 GSPNLYQ-GGRKPWHSINFICAHDGFTLADLVTYNNKNNLANGEENNDGENHNYSWNCGE 574

Qy     559 EGEFARLSVKRLRKRQMRNFFVCLMVSQGVPMFYMGDEYGHTKGGNNNTYCHDSYVNYFR 618
Db     575 EGDFASISVKRLRKRQMRNFFVSLMVSQGVPMIYMGDEYGHTKGGNNNTYCHDNYMNYFR 634

Qy     619 WDKKEQ-YSELHRFCCLMTKFRKECEGLGLEDFPTAKRLQWHGHQPGKPDWSENSRFVAF 677
Db     635 WDKKEEAHSDFFRFCRILIKFRDECESLGLNDFPTAKRLQWHGLAPEIPNWSETSRFVAF 694

Qy     678 SMKDERQGEIYVAFNTSHLPAVVELPERAGRRWEPVVDTGKPAPYDFLTDDLPDRALTIH 737
Db     695 SLVDSVKKEIYVAFNTSHLATLVSLPNRPGYRWEPFVDTSKPSPYDCITPDLPERETAMK 754

Qy     738 QFSHFLYSNLYPMLSYSSVILVLRP 762
Db     755 QYRHFLDANVYPMLSYSSIILLLSP 779
```

RESULT 4
S77094
glycogen operon protein (EC 3.2.1.-) glgX-2 - Synechocystis sp. (strain PCC 6803)
N;Alternate names: protein slr1857
C;Species: Synechocystis sp.
A;Variety: PCC 6803
C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
C;Accession: S77094
R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
o, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A;Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
s.
A;Reference number: S74322; MUID:97061201
A;Accession: S77094
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-707 <KAN>
A;Cross-references: EMBL:D90908; GB:AB001339; NID:g1652725; PIDN:BAA17652.1; PID:g165273
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

C;Genetics:
A;Gene: glgX-2
C;Superfamily: glyX protein
C;Keywords: glycosidase; hydrolase

Query Match 38.9%; Score 1638.5; DB 2; Length 707;
Best Local Similarity 49.1%; Pred. No. 5.6e-118;
Matches 342; Conservative 77; Mismatches 216; Indels 61; Gaps 16;

```
Qy      56 RVLAGMPAPLGATALAGGVNFAVYSGGATAAALCLFTPEDLKADRVTEEVPLDPLMNRTG 115
Db      16 KLRCGQPFPFGATIVPGGVNFSIYSSHSTACTLVLF---EKRAPQPFVEIPF-PESFRIG 71

Qy     116 NVW-HVFIEGELHNMLYGYRFDGTFAPHCGHYLDVSNVVVDPYAKAVISRGEYGVPARGN 174
Db      72 NVYCMVVFDLDFENLEYGYRMEGPNNFQQGHWFDPSKVLLDPYAKVVSGRDVWGTQPNWD 131

Qy     175 NCWPQMAGMIPLPYSTFDWEGDLPLRYPQKDLVIYEMHLRGFTKHDSSNVE--HPGTFIG 232
Db     132 DIY-QHRGR--LSFDDFDWENDSPLDVPLEDMVIYEMHVRGFTKDPSSGVKENHRGTFAG 188

Qy     233 AVSKLDYLKELGVNCIELMPCHEFNELEYS-----TSSSKMNFWGYSTINFFSPMTRYTS 287
Db     189 ILSKIPYLQELGVNTIELMPIFEFDEFEHSRYHPETGEFLVNYWGYSTVNFFAPKAGYAA 248

Qy     288 GGIKNCGRDAINEFKTFVREAHKRGIEVILDVVFNHTAEGNENGPILSFKGVDNTTYYML 347
Db     249 TG--KFGMQ-IDELKNLVKELHKVGISVILDVVFNHTAEGNERGPTISFRGLDNKTYYML 305

Qy     348 APKGEFYNYSGCGNTFNCNHPVVRQFIVDCLRYWVTEMHVDGFRFDLASIMTRGSSLWDP 407
Db     306 TPEGYYFNFSGTGNTLNCNNPIVRGMVLDCLRYWTAEFHIDGFRFDLASILGR-----DP 360

Qy     408 VNVYGAPIEGDMITTGTPLVTPPLIDMISNDPILGGVKLIAEAWDAGGLYQVGQFPHWNV 467
Db     361 --------------WGYPLANPPLLETLAFDPILARSKLIAEAWDAGGLYQVGSFPSYGR 406

Qy     468 WSEWNGKYRDIVRQFIKGTDGFAGGFAECLCGSPHLYQAGGRKPWHSINFVCAHDGFTLA 527
Db     407 WAEWNGKYRDTVRKFIKGDAGVIGEMAQRLQGSPDLYQGAGRPPSTSINFVTAHDGFTLA 466

Qy     528 DLVTYNKKYNLPNGENNRDGENHNLSWNCGEEGEFARLSVKRLRKRQMRNFFVCLMVSQG 587
Db     467 DLVAYNGKHNYANGENGNDGANDNYSWNCGVEGPTDNPDILRLRARQMRNAIAILLVSQG 526

Qy     588 VPMFYMGDEYGHTKGGNNNTYCHDSYVNYFRWDKKEQYSELHRFCCLMTKFR-------- 639
Db     527 VPMLLMGDEMGKTQDGNNNTYCHDSPFNWLNWHLLEQNQAWFRFVKHCIAFRLAHPVLRN 586

Qy     640 ----KECEGLGLEDFPTAKRLQWHGHQPGKPDWSENSRFVAFSMKDE-------RQGEIY 688
Db     587 SEHFQNCDYLGV-GFPD---ISWHGVKPWHADWSADSRVLAFMLCGRHAKGGRVKDNQIY 642

Qy     689 VAFNTSHLPAVVELP-ERAGRRWEPVVDTGKPAPYD 723
Db     643 VAMNMHYESLWFELPAPPVGTTWHVFANTGAQPPED 678
```

RESULT 5
B85091
isoamylase-like protein [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 02-Mar-2001
C;Accession: B85091
R;anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold S
Nature 402, 769-777, 1999
A;Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A;Reference number: A85001; MUID:20083488
A;Accession: B85091
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-702 <STO>

```
A;Cross-references: GB:NC_001268; NID:g7267544; PIDN:CAB78026.1; GSPDB:GN00140
C;Genetics:
A;Gene: AT4g09020
A;Map position: 4
C;Superfamily: glyX protein


  Query Match             36.5%;  Score 1538.5;  DB 2;  Length 702;
  Best Local Similarity   44.0%;  Pred. No. 2.8e-110;
  Matches  325;  Conservative  110;  Mismatches  222;  Indels   81;  Gaps   19;

Qy      56 RVLAGMPAPLGATALAGGVNFAVYSGGATAAALCLFT---------PEDLKADRVTEEVP 106
           :| :|  :||| : :  |:|||::|  ||:  |||            :|   | : | |
Db       8 KVSSGEVSPLGVSQVDKGINFALFSQNATSVTLCLSLSQRYITSSGKDDTDDDGMIELV- 66

Qy     107 LDPLMNRTGNVWHVFIEG-ELHNMLYGYRFDGTFAPHCGHYLDVSNVVVDPYAKAVISRG 165
           ||| :|:||: ||: :|   |:|:|||| ||      ||  | | :::|||| |
Db      67 LDPSVNKTGDTWHICVEDLPLNNVLYGYRVDGPGEWQQGHRFDRSILLLDPYAKLVKGHS 126

Qy     166 EYGVPARGNNCWPQMAGMIPLPYSTFDWEGDLPL-RYPQKDLVIYEMHLRGFTKHDSSNV 224
            :|     :  : |  |      | |||  |      |:||||||||::| ||  :|| :
Db     127 SFG---DSSQKFAQFYGTYDFESSPFDWGDDYKFPNIPEKDLVIYEMNVRAFTADESSGM 183

Qy     225 EHP--GTFIGAVSKLDYLKELGVNCIELMPCHEFNELEYSTSSS----KMNFWGYSTINF 278
           :    |:::| : |: :|::||:| :||:|  ||:|||    |:     :| ||||||:||
Db     184 DPAIGGSYLGFIEKIPHLQDLGINAVELLPVFEFDELELQRRSNPRDHMVNTWGYSTVNF 243

Qy     279 FSPMTRYTSG---GIKNCGRDAINEFKTFVREAHKRGIE---------VILDVVFNHTAE 326
           |:||:|| ||    ||      |  |||  |:  |  |||          ||||||:||| |
Db     244 FAPMSRYASGEGDPIK-----ASKEFKEMVKALHSAGIEKYSYKFSLQVILDVVYNHTNE 298

Qy     327 GNENGP-ILSFKGVDNTTYYMLAPKGEFYNYSGCGNTFNCNHPVVRQFIVDCLRYWVTEM 385
            ::  |   ||:|:||  |||| |  :  |:|||||| |||||||  : |:|  ||:|||
Db     299 ADDKYPYTTSFRGIDNKVYYMLDPNNQLLNFSGCGNTLNCNHPVVMELILDSLRHWVTEY 358

Qy     386 HVDGFRFDLASIMTRGSSLWDPVNVYGAPIEGDMITTGTPLVTPPLIDMISNDPILGGVK 445
           ||||||||||:: |                  | |:||  ||||  |: | :|   |
Db     359 HVDGFRFDLASVLCRA------------------TDGSPLSAPPLIRAIAKDSVLSRCK 399

Qy     446 LIAEAWDAGGLYQVGQFPHWNVWSEWNGKYRDIVRQFIKGTDGFAGGFAECLCGSPHLYQ 505
           :||| || |||| ||:||:|: |:|||| ||| ||:|||   |  | ||  :  ||  |||
Db     400 IIAEPWDCGGLYLVGKFPNWDRWAEWNGMYRDDVRRFIKGDSGMKGSFATRVSGSSDLYQ 459

Qy     506 AGGRKPWHSINFVCAHDGFTLADLVTYNKKYNLPNGENNRDGENHNLSWNCGEEGEFARL 565
              |||:| :||| ||||||| |||:|| |:|  |||    || | | |||||  |||
Db     460 VNQRKPYHGVNFVIAHDGFTLRDLVSYNFKHNEANGEGGNDGCNDNHSWNCGFEGETGDA 519

Qy     566 SVKRLRKRQMRNFFVCLMVSQGVPMFYMGDEYGHTKGGNNNTYCHDSYVNYFRWDKKEQY 625
           :| || |||:|| : ||:||| ||  |||||||| ||||:| ||: :| |:| : :
Db     520 HIKSLRTRQMKNFHLALMISQGTPMMLMGDEYGHTRYGNNNSYGHDTSLNNFQWKELDAK 579

Qy     626 SELH-RFCCLMTKFRKECEGLGLEDFPTAKRLQWHGHQPGKPDW-SENSRFVAFSMKDER 683
           : | ||   : |||      |  |:| |   : ||      : :| :  |:|:||:: |
Db     580 KQNHFRFFSEVIKFRHSHHVLKHENFLTQGEITWH-----EDNWDNSESKFLAFTLHDGI 634

Qy     684 QG-EIYVAFNTS--HLPAVVELPERAGRRWEPVVDTGKPAPYDFLTDDLPDRALTIHQFS 740
           | :||||||     : |::  |   |::|  | ||   :| ||: : :   | |
Db     635 GGRDIYVAFNAHDYFVKALIPQPP-PGKQWFRVADTNLESPDDFVREGVAGVADT----- 688

Qy     741 HFLYSNLYPMLSYSSVIL 758
                 | :  :||::|
Db     689 -------YNVAPFSSILL 699


RESULT   6
S74546
glycogen operon protein (EC 3.2.1.-) glgX-1 - Synechocystis sp. (strain PCC 6803)
N;Alternate names: protein slr0237
C;Species: Synechocystis sp.
A;Variety: PCC 6803
C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
C;Accession: S74546
R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
o, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Ya
DNA Res. 3, 109-136, 1996
A;Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocy
s.
A;Reference number: S74322; MUID:97061201
A;Accession: S74546
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-746 <KAN>
A;Cross-references: EMBL:D90900; GB:AB001339; NID:g1651768; PIDN:BAA16698.1; PID:g16
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C;Genetics:
A;Gene: glgX-1
A;Start codon: GTG
C;Superfamily: glyX protein
C;Keywords: glycosidase; hydrolase


  Query Match             33.0%;  Score 1388;  DB 2;  Length 746;
  Best Local Similarity   42.0%;  Pred. No. 1.2e-98;
  Matches  305;  Conservative   86;  Mismatches  251;  Indels   84;  Gaps   17;

Qy      60 GMPAPLGATALAGGVNFAVYSGGATAAALCLF-TPEDLKADRVTEEVPLDPLMNRTGNVW 118
           |   |||||  | |||| ::|  |    | ||  | |    |  |   |    |||   |
Db      17 GQTFPLGATVYADGVNFCLFSKHAERVTLLLFDRPNDPAPARTIE---LHRGRNRTFYYW 73

Qy     119 HVFIEGELHNMLYGYRFDGTFAPHCGHYLDVSNVVVDPYAKAVISRGEY---GVPARGNN 175
           |||::|     :| || ||    |  ||  |    |::||||||:: :  |       | |:|
Db      74 HVFVKGLKAGQVYAYRVDGPHEPEKGHRFDPDKVLLDPYAKAIVGKDIYDRKAAMALGDN 133

Qy     176 CWPQMAGMIPLPYSTFDWEGDLPLRYPQKDLVIYEMHLRGFTKHDSSNVEH--PGTFIGA 233
           |  |    : :  | :||| |   | |    :|||:|: |||:: :| :     ||: |
Db     134 C-AQALRSVVVDTSVYDWEDDHAPRTPYAASIIYELHVGGFTRNPNSGLSENKRGTYAGL 192

Qy     234 VSKLDYLKELGVNCIELMPCHEFNELEYSTSSSKMNFWGYSTINFFSPMTRYTSGGIKNC 293
           : |: ||||||:  :||:| | |:  :        |:|||||| ||:|   |::    :
Db     193 IEKIPYLKELGITAVELLPVHYFDPED--AQPGLTNYWGYSTIGFFAPHQGYSAA---DD 247

Qy     294 GRDAINEFKTFVREAHKRGIEVILDVVFNHTAEGNENGPILSFKGVDNTTYYML-APKGE 352
           : ::||:  |:  || ||||||||||||||||| || |||:|:|| |||:|   |
Db     248 PLEVVDEFRDMVKALHKAGIEVILDVVFNHTAEGNEKGPTLSFRGIDNRTYYILDEDKSS 307

Qy     353 FYNYSGCGNTFNCNHPVVRQFIVDCLRYWVTEMHVDGFRFDLASIMTRGSSLWDPVNVYG 412
           : |||||||:   |||||   |:| |||||:|||||||||||||:: | :         |
Db     308 YSNYSGCGNSVKANHPVVGGLILDSLRYWVSEMHVDGFRFDLASVLVRDTK--------G 359

Qy     413 APIEGDMITTGTPLVTPPLIDMISNDPILGGVKLIAEAWDAGGLYQVGQFPHWNVW-SEW 471
           |: |  | |         :|  | :|||| | |||||||| ||| ||:|     | :||
Db     360 VPLHGSEIATAN------IIWAIESDPILAGTKLIAEAWDAAGLYSVGKFVELADWFAEW 413

Qy     472 NGKYRDIVRQFIKGTDGFAGGFAECLCGSPHLYQAGGRKPWHSINFVCAHDGFTLADLVT 531
           || :|| ||:|:|| :|     |  | ||| :|          |||||  |||||| |||:
Db     414 NGPFRDDVRRFVKGDNGAVPALASRLLGSPDIYYRQDTDINRSINFVTCHDGFTLVDLVS 473

Qy     532 YNKKYNLPNGENNRDGENHNLSWNCGEEGEFARLSVKRLRKRQMRNFFVCLMVSQGVPMF 591
           ||:|:|  ||| |||| | | |||||  |||      : :|| ||::||   |  ||| ||
Db     474 YNEKHNEANGEKNRDGTNDNFSWNCGVEGETDDPKINQLRLRQIKNFLTILFFSQGTPML 533

Qy     592 YMGDEYGHTKGGNNNTYCHDSYVNYFRWDKKEQYSELHRFC------------------ 632
           |||  | |: |||| || |: :::| |  :| :|:   |
Db     534 LMGDPVGRTQQGNNNGYCQDNELSWFDWSGEETHSDETHFLRGIIALTQSLSLFLEESLL 593

Qy     633 ----CLMTKFRKE---CEGLGLE-DFPTAKRLQ----------------------WHG 660
                : :|  |     || | : |  | |                        |||
Db     594 PVLNFALPRFPDENDGSEGDGQDGEEPVYKPLPPVDEAALREMYLKATKDLPGSCVVWHG 653

Qy     661 HQPGKPDWSENSRFVAFSMKDERQGEI-YVAFNTSHLPAVVELPE-RAGRRWEPVVDTGK 718
           :  :|||| |  :| ::      || :: ||     |   |||     |  |  :|||
Db     654 VKLAQPDWSYTSHTLAVTLFHPGAEEILHLIFNAYWEPLNFELPPLNDGLIWHRLVDTYL 713

Qy     719 PAPYDF 724
           | | ||
Db     714 PTPLDF 719
```

```
RESULT   7
B86539
glycogen hydrolase [imported] - Chlamydophila pneumoniae (strain J138)
C;Species: Chlamydophila pneumoniae, Chlamydia pneumoniae
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
C;Accession: B86539
R;Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Is
Nucleic Acids Res. 28, 2311-2314, 2000
A;Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
A;Reference number: A86491; MUID:20330349
A;Accession: B86539
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-664 <STO>
A;Cross-references: GB:BA000008; NID:g8978760; PIDN:BAA98596.1; GSPDB:GN00142
A;Experimental source: strain J138
C;Genetics:
A;Gene: glgX
C;Superfamily: glyX protein


  Query Match             31.0%;  Score 1305.5;  DB 2;  Length 664;
  Best Local Similarity   42.6%;  Pred. No. 2.3e-92;
  Matches  279;  Conservative   88;  Mismatches  231;  Indels   57;  Gaps   16;

Qy      61 MPAPLGATALA-GGVNFAVYSGGATAAALCLFTPEDLKADRVTE--EVPLDPLMNRTGNV 117
           :| ||||: ::     ||:|:  ||   | |        |  :|  |||| |  :||| :
Db      10 VPLPLGASKISPNRYRFALYASQATEVILAL-------TDENSEVIEVPLYPDTHRTGAI 62

Qy     118 WHVFIEGELHNMLYGYRFDGTFAPHCGHYLDVSNVVVDPYAKAVISRGEYGVPARGNN-- 175
           ||: |||      | :|  |    | |        : ||||| : |   :|   :  :
Db      63 WHIEIEGISDQSSYAFRVHGP-KKH-GMQYSFKEYLADPYAKNIHSPQSFGSRKKQGDYA 120

Qy     176 -CWPQMAGMIPLPYSTFDWEGDLPLRYPQKDLVIYEMHLRGFTKHDSSNVEHPGTFIGAV 234
            |:        |    | |:|| ||  |:::::|||||:| ||:  || |  ||||:| :
Db     121 FCY--------LKEEPFPWDGDQPLHLPKEEMIIYEMHVRSFTQSSSSRVHAPGTFLGII 172

Qy     235 SKLDYLKELGVNCIELMPCHEFNELEYSTSSSKM----NFWGYSTINFFSPMTRYTSGGI 290
           |:|:| :||:| :||:|  ||:|  :   :||     |:||| :|||||  ||
Db     173 EKIDHLHKLGINAVELLPIFEFDETAHPFRNSKFPYLCNYWGYAPLNFFSPCRRYAYAS- 231

Qy     291 KNCGRDAINEFKTFVREAHKRGIEVILDVVFNHTAEGNENGPILSFKGVDNTTYYMLAPK 350
             |      |||| |:  |: |||||||||||||      |   |   :|  :||:|  :
Db     232 DPCAPS--REFKTLVKTLHQEGIEVILDVVFNHT---GLQGTTCSLPWIDTPSYYILDAQ 286

Qy     351 GEFYNYSGCGNTFNCNHPVVRQFIVDCLRYWVTEMHVDGFRFDLASIMTRGSSLWDPVNV 410
           | | |||||||| | |     |:|:| ||||| ||||||||||||: :|| |
Db     287 GHFTNYSGCGNTLNTNRAPTTQWILDILRYWVEEMHVDGFRFDLASVFSRGPS------- 339

Qy     411 YGAPIEGDMITTGTPLVTPPLIDMISNDPILGGVKLIAEAWDAGGLYQVGQFPHWNV-WS 469
                   |:||   |::: || ||:|   |:||| |||||||||| ||  :  ||
Db     340 -----------GSPLQFAPVLEAISFDPLLASTKIIAEPWDAGGLYQVGYFPTLSPRWS 387

Qy     470 EWNGKYRDIVRQFIKGTDGFAGGFAECLCGSPHLYQAGGRKPWHSINFVCAHDGFTLADL 529
           |||| ||| |: |: |     | ||  : ||  :|  |   | :|||:|  ||||||
Db     388 EWNGPYRDNVKAFLNGDQNLIGTFASRISGSQDIYPHG--SPTNSINYVSCHDGFTLCDT 445

Qy     530 VTYNKKYNLPNGENNRDGENHNLSWNCGEEGEFARLSVKRLRKRQMRNFFVCLMVSQGVP 589
           |||| |:|  |||:|||| : | |:| | ||:      :  :|:||:|||| :||||:|
Db     446 VTYNHKHNEANGEDNRDGTDANYSYNFGTEGKTEDPGILEVRERQLRNFFLTLMVSQGIP 505

Qy     590 MFYMGDEYGHTKGGNNNTYCHDSYVNYFRWDKKEQYSELHRFCCLMTKFRKECEGLGLED 649
           |   |||| ||  |||| :  ||  ||| ||:      |  | | :  |||: : |
Db     506 MIQSGDEYAHTAEGNNNRWALDSNANYFLWDQLTAKPTLMHFLCDLIAFRKKYKTLFNRG 565

Qy     650 FPTAKRLQWHGHQPGKPDWSENSRFVAFSMKDERQGEIYVAFNTSHLPAVVELPE 704
           | : | : |         |   : |:|| :|  :   :||||:       :  ||:
Db     566 FLSNKEISWVDAMGNPMTWRPGN-FLAFKIKSPK-AHVYVAFHVGAQDQLATLPK 618


RESULT   8
B72084
glycosyl hydrolase family protein CP0367 [imported] - Chlamydophila pneumoniae (str
C;Species: Chlamydophila pneumoniae, Chlamydia pneumoniae
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-May-2000
C;Accession: B72084; C81586
R;Kalman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood
Nature Genet. 21, 385-389, 1999
A;Title: Comparative genomes of Clamydia pneumoniae and C. trachomatis.
A;Reference number: A72000; MUID:99206606
A;Accession: B72084
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-664 <ARN>
A;Cross-references: GB:AE001623; GB:AE001363; NID:g4376662; PIDN:AAD18532.1; PID:g4
A;Experimental source: strain CWL029
R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hicl
, C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzl
Nucleic Acids Res. 28, 1397-1406, 2000
A;Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR
A;Reference number: A81500; MUID:20150255
A;Accession: C81586
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-664 <REA>
A;Cross-references: GB:AE002198; GB:AE002161; NID:g7189279; PIDN:AAF38216.1; PID:g7
A;Experimental source: strain AR39, HL cells
C;Genetics:
A;Gene: glgX; CP0367
C;Superfamily: glyX protein


  Query Match             31.0%;  Score 1305.5;  DB 2;  Length 664;
  Best Local Similarity   42.6%;  Pred. No. 2.3e-92;
  Matches  279;  Conservative   88;  Mismatches  231;  Indels   57;  Gaps   16;

Qy      61 MPAPLGATALA-GGVNFAVYSGGATAAALCLFTPEDLKADRVTE--EVPLDPLMNRTGNV 117
           :| ||||: ::     ||:|:  ||   | |        |  :|  |||| |  :||| :
Db      10 VPLPLGASKISPNRYRFALYASQATEVILAL-------TDENSEVIEVPLYPDTHRTGAI 62

Qy     118 WHVFIEGELHNMLYGYRFDGTFAPHCGHYLDVSNVVVDPYAKAVISRGEYGVPARGNN-- 175
           ||: |||      | :|  |    | |        : ||||| : |   :|   :  :
Db      63 WHIEIEGISDQSSYAFRVHGP-KKH-GMQYSFKEYLADPYAKNIHSPQSFGSRKKQGDYA 120

Qy     176 -CWPQMAGMIPLPYSTFDWEGDLPLRYPQKDLVIYEMHLRGFTKHDSSNVEHPGTFIGAV 234
            |:        |    | |:|| ||  |:::::|||||:| ||:  || |  ||||:| :
Db     121 FCY--------LKEEPFPWDGDQPLHLPKEEMIIYEMHVRSFTQSSSSRVHAPGTFLGII 172

Qy     235 SKLDYLKELGVNCIELMPCHEFNELEYSTSSSKM----NFWGYSTINFFSPMTRYTSGGI 290
           |:|:| :||:| :||:|  ||:|  :   :||     |:||| :|||||  ||
Db     173 EKIDHLHKLGINAVELLPIFEFDETAHPFRNSKFPYLCNYWGYAPLNFFSPCRRYAYAS- 231

Qy     291 KNCGRDAINEFKTFVREAHKRGIEVILDVVFNHTAEGNENGPILSFKGVDNTTYYMLAPK 350
             |      |||| |:  |: |||||||||||||      |   |   :|  :||:|  :
Db     232 DPCAPS--REFKTLVKTLHQEGIEVILDVVFNHT---GLQGTTCSLPWIDTPSYYILDAQ 286

Qy     351 GEFYNYSGCGNTFNCNHPVVRQFIVDCLRYWVTEMHVDGFRFDLASIMTRGSSLWDPVNV 410
           | | |||||||| | |     |:|:| ||||| ||||||||||||: :|| |
Db     287 GHFTNYSGCGNTLNTNRAPTTQWILDILRYWVEEMHVDGFRFDLASVFSRGPS------- 339

Qy     411 YGAPIEGDMITTGTPLVTPPLIDMISNDPILGGVKLIAEAWDAGGLYQVGQFPHWNV-WS 469
                   |:||   |::: || ||:|   |:||| |||||||||| ||  :  ||
Db     340 -----------GSPLQFAPVLEAISFDPLLASTKIIAEPWDAGGLYQVGYFPTLSPRWS 387

Qy     470 EWNGKYRDIVRQFIKGTDGFAGGFAECLCGSPHLYQAGGRKPWHSINFVCAHDGFTLADL 529
           |||| ||| |: |: |     | ||  : ||  :|  |   | :|||:|  ||||||
Db     388 EWNGPYRDNVKAFLNGDQNLIGTFASRISGSQDIYPHG--SPTNSINYVSCHDGFTLCDT 445

Qy     530 VTYNKKYNLPNGENNRDGENHNLSWNCGEEGEFARLSVKRLRKRQMRNFFVCLMVSQGVP 589
           |||| |:|  |||:|||| : | |:| | ||:      :  :|:||:|||| :||||:|
Db     446 VTYNHKHNEANGEDNRDGTDANYSYNFGTEGKTEDPGILEVRERQLRNFFLTLMVSQGIP 505

Qy     590 MFYMGDEYGHTKGGNNNTYCHDSYVNYFRWDKKEQYSELHRFCCLMTKFRKECEGLGLED 649
           |   |||| ||  |||| :  ||  ||| ||:      |  | | :  |||: : |
Db     506 MIQSGDEYAHTAEGNNNRWALDSNANYFLWDQLTAKPTLMHFLCDLIAFRKKYKTLFNRG 565

Qy     650 FPTAKRLQWHGHQPGKPDWSENSRFVAFSMKDERQGEIYVAFNTSHLPAVVELPE 704
           | : | : |         |   : |:|| :|  :   :||||:       :  ||:
Db     566 FLSNKEISWVDAMGNPMTWRPGN-FLAFKIKSPK-AHVYVAFHVGAQDQLATLPK 618
```

```
Db    506 MIQSGDEYAHTAEGNNNRWALDSNANYFLWDQLTAKPTLMHFLCDLIAFRKKYKTLFNRG 565

Qy    650 FPTAKRLQWHGHQPGKPDWSENSRFVAFSMKDERQGEIYVAFNTSHLPAVVELPE 704
Db    566 FLSNKEISWVDAMGNPMTWRPGN-FLAFKIKSPK-AHVYVAFHVGAQDQLATLPK 618


RESULT   9
E71565
probable glycogen hydrolase (debranching) - Chlamydia trachomatis (serotype D, strain UW
C;Species: Chlamydia trachomatis
C;Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 08-Oct-1999
C;Accession: E71565
R;Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell,
Science 282, 754-759, 1998
A;Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trac
A;Reference number: A71570; MUID:99000809
A;Accession: E71565
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-666 <ARN>
A;Cross-references: GB:AE001278; GB:AE001273; NID:g3328423; PIDN:AAC67632.1; PID:g332843
A;Experimental source: serotype D, strain UW-3/Cx
C;Genetics:
A;Gene: glgX
C;Superfamily: glyX protein


  Query Match             30.9%;  Score 1300;  DB 2;  Length 666;
  Best Local Similarity   39.8%;  Pred. No. 6.2e-92;
  Matches  289;  Conservative  102;  Mismatches  241;  Indels   94;  Gaps   20;

Qy     57 VLAGMPAPLGATAL-AGGVNFAVYSGGATAAALCLFTPEDLKADRVTE--EVPLDPLMNR 113
Db      6 VRSTIPLPLGAKKLSADRYRFSLFSSQAQQVTLVLLDP-------LSEIHEIPLSSTDHR 58

Qy    114 TGNVWHVFIEGELHNMLYGYRFDGT------FAPHCGHYLDVSNVVVDPYAKAVISRGEY 167
Db     59 TGAIWHIEIAGISSEWSYAYKLRGTDLSSQKFA--------TDSYIADPYSKNIYSPQLF 110

Qy    168 GVPARGNNCWPQMAGMIPLPYSTFDWEGDLPLRYPQKDLVIYEMHLRGFTKHDSSNVEHP 227
Db    111 GSPKQEKD-----YAFSYLKHEDFDWEGDTPLHLPKENYFIYEMHVRSFTRDPSSQVSHP 165

Qy    228 GTFIGAVSKLDYLKELGVNCIELMPCHEFNELEYSTSSSKM----NFWGYSTINFFSPMT 283
Db    166 GTFLGIIEKIDHLKQLGVHAVELLPIFEFDETVHPFKNQDFPHLCNYWGYSSVNFFCPSR 225

Qy    284 RYTSGGIKNCGRDAINEFKTFVREAHKRGIEVILDVVFNHTAEGNENGPILSFKGVDNTT 343
Db    226 RYTYGA-DPCA--PAREFKTLVKALHRAGIEVILDVVFNHTGFEGTSCPL---PWIDLES 279

Qy    344 YYMLAPKGEFYNYSGCGNTFNCNHPVVRQFIVDCLRYWVTEMHVDGFRFDLASIMTRGSS 403
Db    280 YYMVNDHGDLMNFSGCGNTVNTNTPTTLKWILDALRYWVQEMHVDGFRFDLASVFSR--- 336

Qy    404 LWDPVNVYGAPIEGDMITTGTPLVTPPLIDMISNDPILGGVKLIAEAWDAGGLYQVGQFP 463
Db    337 --DP--------------QGVPLPLTPILQAISSDSILSETKLIAEPWDAGGLYQLGHFP 380

Qy    464 HWNV-WSEWNGKYRDIVRQFIKGTDGFAGGFAECLCGSPHLYQAGGRKPWHSINFVCAHD 522
Db    381 SISTRWSEWNGCYRDHVKAFLNGDAHQVSSFASRISGSHDIYPNG--KPTNSINYICSHD 438

Qy    523 GFTLADLVTYNKKYNLPNGENNRDGENHNLSWNCGEEGEFARLSVKRLRKRQMRNFFVCL 582
Db    439 GFTLYDTVAYNDKHNEENGEYNRDGTSANYSYNFGCEGETTDPTICALRERQMKNFFLAL 498

Qy    583 MVSQGVPMFYMGDEYGHTKGGNNNTYCHDSYVNYFRWDKKEQYSELHRFCCLMTKFRKEC 642
Db    499 FLSQGIPMIQSGDEYGHTAYGNNNHWCLDTKINYFLWDRLAERKELFSFLCQVIALRKAY 558

Qy    643 EGLGLEDFPTAKRLQWHGHQPGKPDWSENSRFVAFSMKDERQGEIYVAFNTSHLPAVVEL 702
```

```
Db    559 TELFNTSFLSEDTITWLNTKGSPREWGAD-HYLAFELK-HLNYSLFVAFYSGNERIEISL 616

Qy    703 PE--RAGRRWEPVVD--TGKPAPYDFLTDDLPDRALTIHQFSHFLYSNLYPMLS---YSS 755
Db    617 PKPRKEHLAYEKIVDSTTG----------------------FFSQILSPKLSLEPYSS 652

Qy    756 VILVLR 761
Db    653 LVAISR 658


RESULT  10
B75540
glycogen operon protein GlgX - Deinococcus radiodurans (strain R1)
C;Species: Deinococcus radiodurans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
C;Accession: B75540
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.
, M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.;
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A;Reference number: A75250; MUID:20036896
A;Accession: B75540
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-720 <WHI>
A;Cross-references: GB:AE001888; GB:AE000513; NID:g6457936; PIDN:AAF09848.1; PID:g645
A;Experimental source: strain R1
C;Genetics:
A;Gene: DR0264
A;Map position: 1
C;Superfamily: glyX protein


  Query Match             30.1%;  Score 1268;  DB 2;  Length 720;
  Best Local Similarity   39.8%;  Pred. No. 2e-89;
  Matches  282;  Conservative   92;  Mismatches  241;  Indels   94;  Gaps   17;

Qy     60 GMPAPLGATALAGGVNFAVYSGGATAAALCLFTPEDLKADRVTEEVPLDPLMNRTGNVWH 119
Db     11 GSPFPLGATWDGKGTNFALYSENATGVELCLFDAEGHET--------RFPLTEQTAFVWH 62

Qy    120 VFIEGELHNMLYGYRFDGTFAPHCGHYLDVSNVVVDPYAKAV-----ISRGEYGVPARGN 174
Db     63 GYLPGIQPGQRYGYRVHGEYAPEKGLRFNPNVVLLDPYAKALDGTEQFDRGVFGYVAGGE 122

Qy    175 NCWPQM---------AGMIPLPYSTFDWEGDLPLRYPQKDLVIYEMHLRGFT-KHDSSNV 224
Db    123 DD-SQMQEEEQRGAPLGLVVDP--MFNWVGDQKPGIPFHQSVIYEAHVKGLTMTHPDVPE 179

Qy    225 EHPGTFIGAVSK--LDYLKELGVNCIELMPCHEFNELEYSTSSSKMNFWGYSTINFFSPM 282
Db    180 ELRGTYAGVATPAILDYLRDLGITAIEFLPVHQHVDDPFLLDKGLTNYWGYSTLNFFAPD 239

Qy    283 TRYTSGGIKNCGRDAINEFKTFVREAHKRGIEVILDVVFNHTAEGNENGPILSFKGVDNT 342
Db    240 VRYSAEARKGNPSGAVPEFKNMVRALHDAGIEVILDVVYNHTAEGNHMGPTMSFKGIDNP 299

Qy    343 TYYMLAPKGE--FYNYSGCGNTFNCNHPVVRQFIVDCLRYWVTEMHVDGFRFDLASIMTR 400
Db    300 TYYRLVADDQRFYFDYTGTGNSLNVRHPQTLQLIMDSLRYWVTEMHVDGFRFDLASTLAR 359

Qy    401 GSSLWDPVNVYGAPIEGDMITTGTPLVTPPLIDMISNDPILGGVKLIAEAWDAG-GLYQV 459
Db    360 GLHEVDQLSGF--------------------FTIIHQDPIISQVKLIAEPWDVGEGGYQV 399

Qy    460 GQFPHWNVWSEWNGKYRDIVRQFIKGTDGFAGGFAECLCGSPHLYQAGGRKPWHSINFVC 519
Db    400 GNFP--VNWAEWNGIYRDDMRSFWKGEGGLASEIGYRITGSSDLYEFNGRKPYASINFVT 457

Qy    520 AHDGFTLADLVTYNKKYNLPNGENNRDGENHNLSWNCGEEGEFARLSVKRLRKRQMRNFF 579
```

```
          ||||||| | ||| :|:|  ||| | || |||::|||| ||      : ||| :|||||
Db    458 AHDGFTLRDSVTYEQKHNEANGEGNNDGHNHNITWNCGVEGPTDDPEINRLRGQQMRNFL 517

Qy    580 VCLMVSQGVPMFYMGDEYGHTKGGNNNTYCHDSYVNYFRWDK---------------KEQ 624
            |:: || ||   |||:| |:|||| || |: :::: |:|                 ::
Db    518 ATLLLGQGTPMILGGDEFGRTQGGNNNAYCQDNDISWYDWEKVDEELLAFTRKLIALRKA 577

Qy    625 YSELHRFCCLMTKFRKECEGLGL--EDFPTAKRLQWHGHQPGKPDWSENSRFVAFSM--- 679
          :  |||        ||   |  :  ||      |:: | :    ||: | :  :  |
Db    578 HPSLHR--------RKFFAGRNIRGEDVRDIVWLRFDGAEMSDEDWN-NPQTQSLGMFLA 628

Qy    680 -----------KDERQGEIYVAFNTSHLPAVVELPERAG-RRWEPVVDT 716
                      |      : :  ::|::    ::|:  |   |: ::||
Db    629 GDGLADVDAEGKPLTDDHLLLLLSSSYVDLPFKMPDLGGCGEWDLLLDT 677


RESULT  11
G81717
glycosyl hydrolase family protein TC0312 [imported] - Chlamydia muridarum (strain Nigg)
C;Species: Chlamydia muridarum, Chlamydia trachomatis MoPn
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-May-2000
C;Accession: G81717
R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
, C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000
A;Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A;Reference number: A81500; MUID:20150255
A;Accession: G81717
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-666 <TET>
A;Cross-references: GB:AE002298; GB:AE002160; NID:g7190343; PIDN:AAF39177.1; PID:g719035
A;Experimental source: strain Nigg (MoPn)
C;Genetics:
A;Gene: TC0312
C;Superfamily: glyX protein


  Query Match             30.0%;  Score 1264;  DB 2;  Length 666;
  Best Local Similarity   38.4%;  Pred. No. 3.7e-89;
  Matches  282;  Conservative  105;  Mismatches  245;  Indels  102;  Gaps   20;

Qy     53 GACRVLAGMPAPLGATALAGG-VNFAVYSGGATAAALCLFTPEDLKADRVTEEVPLDPLM 111
          |:  :    | ||||  |:     |:: :  |    | |  |          |:||
Db      2 GSLSIHLTTPLPLGAQKLSSDRYRFSLVAPRAQQVILVLLDPFS-----EIHEIPLSSTD 56

Qy    112 NRTGNVWHVFIEGELHNMLYGYRF---DGTFAPHCGHYLDVSNVVVDPYAKAVISRGEYG 168
          :||| :||: | |   :   | |:    |  | |:       :  : |||:| : |
Db     57 HRTGAIWHIEISGISNEWSYAYKLRQSDSAF-PN----FSTNAYIADPYSKNIFS----- 106

Qy    169 VPARGNNCWPQMAGMIPLP---------YSTFDWEGDLPLRYPQKDLVIYEMHLRGFTKH 219
                ||: |    |            |||||| ||| |:::  |||||:| ||:
Db    107 ---------PQLFGSSKQPNDYTFSYLKQEDFDWEGDTPLRLPKENYFIYEMHVRSFTQD 157

Qy    220 DSSNVEHPGTFIGAVSKLDYLKELGVNCIELMPCHEFNELEYSTSSSKM----NFWGYST 275
           || | |||||:| : |:|:||:|||| :||:|  ||:|  :   :        |:|||:
Db    158 PSSQVTHPGTFLGIIEKIDHLKKLGVNAVELLPIFEFDETIHPFKNKDFPHLCNYWGYSS 217

Qy    276 INFFSPMTRYTSGGIKNCGRDAINEFKTFVREAHKRGIEVILDVVFNHTAEGNENGPILS 335
          |||| |  |||  |   |      |||| |:  |: ||||||||||||     : |:
Db    218 INFFCPSRRYTY-GTDPCA--PAREFKTLVKTLHRAGIEVILDVVFNHTGFEGTSCPL-- 272

Qy    336 FKGVDNTTYYMLAPKGEFYNYSGCGNTFNCNHPVVRQFIVDCLRYWVTEMHVDGFRFDLA 395
             :|  :|||:   |:  |:|||||| | | |:  ::|:| ||||| ||||||||||
Db    273 -PWIDLESYYMVNDNGDLLNFSGCGNTVNTNTPIAIKWILDALRYWVQEMHVDGFRFDLA 331

Qy    396 SIMTRGSSLWDPVNVYGAPIEGDMITTGTPLVTPPLIDMISNDPILGGVKLIAEAWDAGG 455
          :: :|                | |    |::  ||:| ||   ||||| |||||
Db    332 AVFSRD------------------LQGVPRSLTPILQAISSDSILSETKLIAEPWDAGG 372

Qy    456 LYQVGQFPHWNV-WSEWNGKYRDIVRQFIKGTDGFAGGFAECLCGSPHLYQAGGRKPWHS 514
          |||:| ||  :  ||||||  ||| |: |: |       ||  : ||  :| ||     :|
Db    373 LYQLGHFPSISTRWSEWNGCYRDHVKAFLNGDPHQVSSFASRISGSRDIYPAGNST--NS 430
```

```
Qy    515 INFVCAHDGFTLADLVTYNKKYNLPNGENNRDGENHNLSWNCGEEGEFARLSVKRLRKRQ 574
          ||::|:|||||| | | || |:|  |||||||| : | |:| | |||    :: :||:||
Db    431 INYICSHDGFTLYDSVAYNHKHNEENGENNRDGTSANYSYNFGCEGETTDPNICQLRERQ 490

Qy    575 MRNFFVCLMVSQGVPMFYMGDEYGHTKGGNNNTYCHDSYVNYFRWDKKEQYSELHRFCCL 634
          |:|||: | :|||:||   |||||||  |||| :| |: :|:| ||:  :  |   | |
Db    491 MKNFFLALFLSQGIPMIKSGDEYGHTAYGNNNHWCLDTKINHFLWDRLAERKEFFSFLCQ 550

Qy    635 MTKFRKECEGLGLEDFPTAKRLQWHGHQPGKPDWSENSRFVAFSMKDERQGEIYVAFNTS 694
          :   |     |   :| : : : |   |    :|: :  ::|| :|      :::|| :
Db    551 IITLRTTHAELFNTNFLSEETITWLNSQGLPREWTPD-HYLAFELKHPNY-SLFIAFYSG 608

Qy    695 HLPAVVELPE--RAGRRWEPVVD--TGKPAPYDFLTDDLPDRALTIHQFSHFLYSNLYPM 750
          :    : ||:  :    :| :||  ||                        |   | |
Db    609 NERIEIALPKLRQEHLAYEKIVDSTTG----------------------FFSQILSPK 644

Qy    751 LS---YSSVILVLR 761
          ||   |||:: : |
Db    645 LSLEPYSSLVAISR 658


RESULT  12
A70764
probable glgX protein - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C;Accession: A70764
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordo
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroy
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Sqares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete ge
A;Reference number: A70500; MUID:98295987
A;Accession: A70764
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-721 <COL>
A;Cross-references: GB:Z74020; GB:AL123456; NID:g3261584; PIDN:CAA98327.1; PID:g140
A;Experimental source: strain H37Rv
C;Genetics:
A;Gene: glgX
C;Superfamily: glyX protein


  Query Match             29.1%;  Score 1225;  DB 2;  Length 721;
  Best Local Similarity   39.4%;  Pred. No. 4.2e-86;
  Matches  280;  Conservative   85;  Mismatches  263;  Indels   82;  Gaps   19;

Qy     57 VLAGMPAPLGATALAGGVNFAVYSGGATAAALCLFTPEDLKADRVTEEVPLDPLMNRTGN 116
          |  |   |||||    | ||:::|  |    |||     :  | |   :||| :    |
Db     19 VWPGNAYPLGATYDGAGTNFSLFSEIAEKVELCL-----IDEDGVESRIPLDEV---DGY 70

Qy    117 VWHVFIEGELHNMLYGYRFDGTFAPHCGHYLDVSNVVVDPYAKAVISRGEYG-------- 168
          ||| ::       ||:|  | | |  ||  | | :::||| |:      :|
Db     71 VWHAYLPNITPGQRYGFRVHGPFDPAAGHRCDPSKLLLDPYGKSFHGDFTFGQALYSYDV 130

Qy    169 -------VPARGNNCWPQMAGMIPLPYSTFDWEGDLPLRYPQKDLVIYEMHLRGFTK-HD 220
                |   ::    |  ::  |:  |||  |   | |  : |||| |::| |: |
Db    131 NAVDPDSTPPMVDSLGHTMTSVVINPF--FDWAYDRSPRTPYHETVIYEAHVKGMTQTHP 188

Qy    221 SSNVEHPGTFIGAVSK--LDYLKELGVNCIELMPCHEFNELEYSTSSSKMNFWGYSTINF 278
          |   |  ||: |       :|:| || |  :|||| |:|            |:|||:|  |
Db    189 SIPPELRGTYAGLAHPVIIDHLNELNVTAVELMPVHQFLHDSRLLDLGLRNYWGYNTFGF 248

Qy    279 FSPMTRYTSGGIKNCGRDAINEFKTFVREAHKRGIEVILDVVFNHTAEGNENGPILSFKG 338
          |:|  :| |   :  |  |: |||| ||  |: |||||||||:|||||||  || ::|:|
Db    249 FAPHHQYAS--TRQAG-SAVAEFKTMVRSLHEAGIEVILDVVYNHTAEGNHLGPTINFRG 305

Qy    339 VDNTTYYMLAPKG-EFY-NYSGCGNTFNCNHPVVRQFIVDCLRYWVTEMHVDGFRFDLAS 396
          :||| || |      || :::| ||: |  ||   | |:| ||||| ||||||||||||
Db    306 IDNTAYYRLMDHDLRFYKDFTGTGNSLNARHPHTLQLIMDSLRYWVIEMHVDGFRFDLAS 365
```

```
Qy    397 IMTRGSSLWDPVNVYGAPIEGDMITTGTPLVTPPLIDMISNDPILGGVKLIAEAWDAG-G 455
Db    366 TLARELHDVDRLSAF-------------------FDLVQQDPVVSQVKLIAEPWDVGEG 405

Qy    456 LYQVGQFPHWNVWSEWNGKYRDIVRQFIKGTDGFAGGFAECLCGSPHLYQAGGRKPWHSI 515
Db    406 GYQVGNFP--GLWTEWNGKYRDTVRDYWRGEPATLGEFASRLTGSSDLYEATGRRPSASI 463

Qy    516 NFVCAHDGFTLADLVTYNKKYNLPNGENNRDGENHNLSWNCGEEGEFARLSVKRLRKRQM 575
Db    464 NFVTAHDGFTLNDLVSYNDKHNEANGENNRDGESYNRSWNCGVEGPTDDPDILALRARQM 523

Qy    576 RNFFVCLMVSQGVPMFYMGDEYGHTKGGNNNTYCHDSYVNYFRWDKKEQYSELHRFCCLM 635
Db    524 RNMWATLMVSQGTPMIAHGDEIGRTQYGNNNVYCQDSELSWMDWSLVDKNADLLAFARKA 583

Qy    636 TKFRKE---------CEGLGLEDFPTAKRLQW---HGHQPGKPDWSEN-SRFVAFSMKDE 682
Db    584 TTLRKNHKVFRRRRFFEGEPIRSGDEVRDIAWLTPSGREMTHEDWGRGFDRCVAVFLNGE 643

Qy    683 ------RQGE------IYVAFNTSHLPAVVELP-ERAGRRWEPVVDTGKP 719
Db    644 AITAPDARGERVVDDSFLLCFNAHDHDVEFVMPHDGYAQQWTGELDTNDP 693
```

RESULT 13
S73088
glycogen operon protein glgX (EC 3.2.1.-) - Sulfolobus solfataricus
N;Alternate names: protein c0621
C;Species: Sulfolobus solfataricus
C;Date: 09-Oct-1997 #sequence_revision 24-Oct-1997 #text_change 20-Jun-2000
C;Accession: S73088
R;Sensen, C.W.; Klenk, H.P.; Singh, R.K.; Allard, G.; Chan, C.C.Y.; Liu, Q.Y.; Penny, S.
Mol. Microbiol. 22, 175-191, 1996
A;Title: Organizational characteristics and information content of an archaeal genome: 1
A;Reference number: S73076; MUID:97055432
A;Accession: S73088
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-718 <SEN>
A;Cross-references: EMBL:Y08256; NID:g1707679; PIDN:CAA69504.1; PID:g1707700
A;Experimental source: strain P2
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1996
C;Superfamily: glyX protein
C;Keywords: glycosidase; hydrolase

Query Match 28.8%; Score 1211; DB 2; Length 718;
Best Local Similarity 38.2%; Pred. No. 5e-85;
Matches 274; Conservative 101; Mismatches 236; Indels 106; Gaps 19;

```
Qy     60 GMPAPLGATALA--GGVNFAVYSGGATAAALCLFTPEDLKADRVTEEVPLDPLMNRTGNV 117
Db     15 GDPYPLGSNWIEDDDGVNFSLFSENAEKVELLLYSLTNQKYPKEIIEV-----KNKTGDI 69

Qy    118 WHVFIEGELHNMLYGYRFDGTFAPHCGHYLDVSNVVVDPYAKAV---------------- 161
Db     70 WHVFVPGLRPGQLYAYRVYGPYKPELGLRFNPNKVLIDPYAKAINGSVIWNDAVFGYKIG 129

Qy    162 ----------ISRGEYGVPARGNNCWPQMAGMIPLPYSTFDWEGD---LPLRYPQKDLVI 208
Db    130 DQNQDLTYDERDSGEY-VP----------KSVVINPY--FEWDDEDFIKGKKVPLKDTVI 176

Qy    209 YEMHLRGFTKHDSSNVEH-PGTFIGAVSK--LDYLKELGVNCIELMPCHEFNELEYSTSS 265
Db    177 YEVHVKGFTKLRLDLPENIRGTYEGLASEQMISYLKDLGITTVELMPVFHFIDQRFLTDK 236

Qy    266 SKMNFWGYSTINFFSPMTRYTSGGIKNCGRDAINEFKTFVREAHKRGIEVILDVVFNHTA 325
Db    237 GLTNYWGYDPINFFSPECRYSSTG---CLGGQVLSFKKMVNELHNAGIEVIIDVVYNHTA 293

Qy    326 EGNENGPILSFKGVDNTTYYMLAPKGEFY--NYSGCGNTFNCNHPVVRQFIVDCLRYWVT 383
Db    294 EGNHLGPTLSFRGIDNTAYYMLQPDNKRYYLDFTGTGNTLNLSHPRVIQMVLDSLRYWVT 353

Qy    384 EMHVDGFRFDLASIMTRGSSLWDPVNVYGAPIEGDMITTGTPLVTPPLIDMISNDPILGG 443
Db    354 EMHVDGFRFDLAAALAR--------ELYSV----NMLNT--------FFIALQQDPILSQ 393

Qy    444 VKLIAEAWDAG-GLYQVGQFPHWNVWSEWNGKYRDIVRQFIKGTDGFAGGFAECLCGSPH 502
Db    394 VKLIAEPWDVGQGGYQVGNFPY--QWAEWNGKYRDSIRRFWRGEALPYSEIANRLLGSPD 451

Qy    503 LYQAGGRKPWHSINFVCAHDGFTLADLVTYNKKYNLPNGENNRDGENHNLSWNCGEEGEF 562
Db    452 IYLGNNKTPFASINYVTSHDGFTLEDLVSYNQKHNEANGFNNQDGMNENYSWNCGAEGPT 511

Qy    563 ARLSVKRLRKRQMRNFFVCLMVSQGVPMFYMGDEYGHTKGGNNNTYCHDSYVNYFRWDKK 622
Db    512 NDQNVVICREKQKRNFMITLLVSQGTPMILGGDELSRTQRGNNNAFCQDNEITWFDWNLD 571

Qy    623 EQYSELHRFCCLMTKFRKECEGLGLEDFPTAKRL----------QWHGHQPGKPDWSEN 671
Db    572 ERKSKFLEFVKKMIQFYRAHPAFRRERYFQGKKLFGMPLKDVTFYTLEGREVDEKTWSSP 631

Qy    672 SRFVAFSMKDERQGEI-----------YVAFNTSHLPAVVELPERAGRRWEPVVDT 716
Db    632 TQLVIFVLEGSVMDEINMYGERIADDSFLIILNANPNNVKVKFPK---GKWELVISS 685
```

RESULT 14
C83375
probable glycosyl hydrolase PA2160 [imported] - Pseudomonas aeruginosa (strain PAO1)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 24-May-2001
C;Accession: C83375
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.;
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic p
A;Reference number: A82950; MUID:20437337
A;Accession: C83375
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-716 <STO>
A;Cross-references: GB:AE004643; GB:AE004091; NID:g9948178; PIDN:AAG05548.1; GSPDB:G
A;Experimental source: strain PAO1
C;Genetics:
A;Gene: PA2160
C;Superfamily: glyX protein

Query Match 27.3%; Score 1149.5; DB 2; Length 716;
Best Local Similarity 36.4%; Pred. No. 2.7e-80;
Matches 275; Conservative 103; Mismatches 260; Indels 117; Gaps 24;

```
Qy     56 RVLAGMPAPLGATALAGGVNFAVYSGGATAAALCLFTPEDLKADRVTEEVPLDPLMNRTG 115
Db     11 RISEGHPFPLGATWDGLGVNFALFSAHATKVELCLF---DARGEKEIERIELP---EYTD 64

Qy    116 NVWHVFIEGELHNMLYGYRFDGTFAPHCGHYLDVSNVVVDPYAKAVISRGEYGVPARGNN 175
Db     65 EIWHGYLPDAHPGQIYGYRVHGPYEPDAGHRFNPNKLLLDPYAKQLVGRLRWSEALFGYT 124

Qy    176 CWPQMAGM-------IP-LPYS-----TFDWEGDLPLRYPQKDLVIYEMHLRGFTKHDSS 222
Db    125 IGSADADLSFDERDSAPFVPKSKVIDPAFTWAERPPVRVPWDRTVIYEAHLRGL------ 178

Qy    223 NVEHP-------GTFIGAVSK--LDYLKELGVNCIELMPCHEFNELEYSTSSSKMNFWGY 273
Db    179 SMRHPQVPEAVRGTFAGLMNADLLAHIRRLGVTSVELLPIHGFVDDKHLLENGMSNYWGY 238

Qy    274 STINFFSPMTRYTSGGIKNCGRDAINEFKTFVREAHKRGIEVILDVVFNHTAEGNENGPI 333
```

```
Db    239 NSIAFFAPHPAYLASG-------QVNEFKEMVAHLHDAGLELILDVVYNHTAEGNELGPT 291

Qy    334 LSFKGVDNTTYYMLAPKGEFY--NYSGCGNTFNCNHPVVRQFIVDCLRYWVTEMHVDGFR 391
Db    292 LCMRGIDNASYYRLMPDQRRYYINDSGTGNTLDLSHPCVLQMVTDSLRYWATEMRVDGFR 351

Qy    392 FDLASIMTRGSSLWDPVNVYGAPIEGDMITTGTPLVTPPLIDMISNDPILGGVKLIAEAW 451
Db    352 FDLATILGRHPDGFDERHGF--------------------LVACRQDPVLSKCKLIAEPW 391

Qy    452 DAG-GLYQVGQFPHWNVWSEWNGKYRDIVRQFIKGTDGFAGGFAECLCGSPHLYQAGGRK 510
Db    392 DCGPGGYQVGGFPPG--WAEWNDRFRDCVRAYWRGDDGMLPELARRLTASGDLYDQRGRR 449

Qy    511 PWHSINFVCAHDGFTLADLVTYNKKYNLPNGENNRDGENHNLSWNCGEEGEFARLSVKRL 570
Db    450 PYASVNFVTAHDGFTLRDVVSYDHKHNEANGENNADGSDHNLSWNHGCEGPTDDPEIRAL 509

Qy    571 RKRQMRNFFVCLMVSQGVPMFYMGDEYGHTKGGNNNTYCHDSYVNYFRW---DK------ 621
Db    510 RLRQMRNLLSTLLLSQGTPMLVAGDEFSRTQQGNNNVYCQDNELGWIDWRLDDEGRSLLA 569

Qy    622 --------KEQYSELHRFCCLMTKFRKECEGLGLEDFPTAKRLQW---HGHQPGKPDW-S 669
Db    570 FTQRLLALRQRYPILRRGRFLVGEYN---EALGVKD------VTWLAPGGEEMTEEHWHD 620

Qy    670 ENSRFVAFSMKDE-------RQGE---IYVAFNTSHLPAVVELPERA-GRRWEPVVDTGK 718
Db    621 EHARCLGVLLDGRAQPTGILRSGEDATLLLILNAYHDAVSFRLPEVAEGSGWTCLLDTQR 680

Qy    719 PAPYDFLTDDLPDRALTIHQF----SHFLYSNLYP 749
Db    681 P------EDPLGERYPFASEFLVGGRSFLLFELQP 709


RESULT  15
C96001
probable glycosyl hydrolase protein (EC 3.2.1.-) [imported] - Sinorhizobium meliloti (st
C;Species: Sinorhizobium meliloti
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
C;Accession: C96001
R;Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A;Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endo
A;Reference number: A95842; MUID:21396508; PMID:11481431
A;Accession: C96001
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-698 <KUR>
A;Cross-references: GB:AL591985; PIDN:CAC49675.1; PID:g15141162; GSPDB:GN00167
A;Experimental source: strain 1021, megaplasmid pSymB
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
 L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:
A;Gene: glgX2; SMb21446
A;Genome: plasmid
C;Superfamily: glyX protein
C;Keywords: glycosidase; hydrolase


  Query Match             26.7%;  Score 1123.5;  DB 2;  Length 698;
  Best Local Similarity   38.0%;  Pred. No. 2.6e-78;
  Matches  272;  Conservative   84;  Mismatches  252;  Indels  107;  Gaps   20;

Qy     61 MPAPLGATALAGGVNFAVYSGGATAAALCLFTPEDLKADRVTEEVPLDPLMNRTGNVWHV 120
Db     11 MKPELGAEYTGQGTHFAVFSAHAEKIELCLFSEDG------REETARMPLPKREGDIWSG 64

Qy    121 FIEGELHNMLYGYRFDGTFAPHCGHYLDVSNVVVDPYAKAVIS----------------R 164
Db     65 YIEGIGPGTLYGYRAHGPYDPHNGHRFNPNKLLLDPYAKLVAGDLIWDDALFGYTIGSPE 124

Qy    165 GEYGVPARGNNCWPQMAGMIPLPYSTFDWEGDLPLRYPQKDLVIYEMHLRGFT-KHDSSN 223
Db    125 GDLSFDERDSAPF-MVKGVVQDP--TFDWAGEEAIRRPWTETIIYESHVRGMTMTHPGVP 181

Qy    224 VEHPGTFIGAVSK--LDYLKELGVNCIELMPCHEFNELEYSTSSSKMNFWGYSTINFFSP 281
Db    182 DELRGTFLGMASDPIIEHLTKLGVSAIELMPVQYFLDDRYLVERNLRNYWGYQTLGFFAP 241

Qy    282 MTRYTSGGIKNCGRDAINEFKTFVREAHKRGIEVILDVVFNHTAEGNENGPILSFKGVDN 341
Db    242 HDRYLKG-------DRFTEFKTMVKRFHSVGIEVLMDVVYNHTAEGSERGPTLSFRGLDN 294

Qy    342 TTYYMLAPKGEFYNY--SGCGNTFNCNHPVVRQFIVDCLRYWVTEMHVDGFRFDLASIMT 399
Db    295 FSYYRQSPDKPRHTYDTTGTGNTLNVAHPLVLRMVLDSLRYWVGVMHIDGFRFDLASTLG 354

Qy    400 RGSSLWDPVNVYGAPIEGDMITTGTPLVTPPLIDMISNDPILGGVKLIAEAWDAG-GLYQ 458
Db    355 REYMEFDR--------EGG------------FFDAIRQDPILAGVKLIAEPWDVGEGGYQ 394

Qy    459 VGQFPHWNVWSEWNGKYRDIVRQFIKGTDGFAGGFAECLCGSPHLYQAGGRKPWHSINFV 518
Db    395 LGGFP--PPFREWNDRFRDDVRRFWKGDGGMVPVLAERILGSPVQFRHSDRAATSSINLV 452

Qy    519 CAHDGFTLADLVTYNKKYNLPNGENNRDGENHNLSWNCGEEGEFARLSVKRLRKRQMRNF 578
Db    453 SAHDGFTLMDTVSYGEKHNEANGEENRDGHSDNHSDNMGAEGPTDNREIIEARKRRRAAM 512

Qy    579 FVCLMVSQGVPMFYMGDEYGHTKGGNNNTYCHDSYVNYFRWDKKEQYSELHRFCCLMTKF 638
Db    513 LATLMISQGVPMVLGGDELGNSQQGNNNAYCQDNEIGWLDWSGAD--DGFVDFCRKLLAF 570

Qy    639 RKECEGLGLEDFPTAKRLQWHGHQPGKPDWSENSRF-VAF-----SMKDE---------- 682
Db    571 RKDRPVLRQERF----------LDGEPD--ENGRVEIAWYKADGSPMDEDAWHDAELRL 617

Qy    683 --------------RQGEIYVAFNTSHLPAVVELPERAG-RRWEPVVDTGKPAPY 722
Db    618 ICAYICRAGQGEAPPAGEIVLVLNAGG-DCEIALPHGNGSRRWLRVLDTAADEPF 671


Search completed: July 31, 2002, 12:08:06
Job time: 156 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: July 31, 2002, 12:11:11 ; Search time 25.26 Seconds
(without alignments)
1171.090 Million cell updates/sec

Title: US-09-674-817A-2
Perfect score: 4212
Sequence: 1 SGPAPRLRRWRPNATAGKGV.....SNLYPMLSYSSVILVLRPDV 764

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1225	29.1	721	1	GLGX_MYCTU	Q10767 mycobacteri
2	1014	24.1	659	1	GLGX_HAEIN	P45178 haemophilus
3	1001	23.8	657	1	GLGX_ECOLI	P15067 escherichia
4	810	19.2	777	1	ISOA_FLASP	O32611 flavobacter
5	803.5	19.1	776	1	ISOA_PSESP	P26501 pseudomonas
6	798.5	19.0	776	1	ISOA_PSEAY	P10342 pseudomonas
7	523	12.4	843	1	PULA_THEMA	O33840 thermotoga
8	310.5	7.4	562	1	O16G_BACTR	P29094 bacillus th
9	310	7.4	598	1	TREZ_ARTSP	Q44316 arthrobacte
10	283.5	6.7	596	1	TREZ_RHISP	Q53238 rhizobium s
11	278	6.6	575	1	TREZ_ARTRM	Q9ajn6 arthrobacte
12	277.5	6.6	589	1	TREZ_BREHE	O52520 brevibacter
13	272	6.5	580	1	TREZ_MYCTU	Q10769 mycobacteri
14	265	6.3	1090	1	PULA_KLEPN	P07206 klebsiella
15	255	6.1	585	1	NEPU_THEVU	Q08751 thermoactin
16	249	5.9	588	1	NEPU_BACST	P38940 bacillus st
17	247.5	5.9	639	1	GLGB_BACST	P30538 bacillus st
18	247	5.9	1096	1	PULA_KLEAE	P07811 klebsiella
19	246.5	5.9	591	1	CDAS_BACSH	Q08341 bacillus sp
20	246.5	5.9	1481	1	APU_THEET	P38939 t amylopull
21	245.5	5.8	728	1	GLGB_ECOLI	P07762 escherichia
22	243.5	5.8	666	1	GLGB_BACCL	P30537 bacillus ca
23	241.5	5.7	558	1	O16G_BACCE	P21332 bacillus ce
24	241	5.7	702	1	GLGB_HUMAN	Q04446 homo sapien
25	241	5.7	799	1	GLGB_MAIZE	Q08047 zea mays (m
26	238.5	5.7	586	1	AMYM_BACAD	P32818 bacillus ac
27	236	5.6	561	1	O16G_BACHD	Q9k8u9 bacillus ha
28	234.5	5.6	639	1	GLGB_BUTFI	P30539 butyrivibri
29	231	5.5	535	1	DEXB_STRPN	Q54796 streptococc
30	228.5	5.4	773	1	GLGB_SYNP7	P16954 synechococc
31	228	5.4	498	1	AMY3_DICTH	P14899 dictyoglomu
32	228	5.4	731	1	GLGB_MYCTU	Q10625 mycobacteri
33	221.5	5.3	1475	1	APU_THETY	P16950 t amylopull
34	220	5.2	770	1	GLGB_SYNY3	P52981 synechocyst
35	219.5	5.2	567	1	MAL1_APIME	Q17058 apis mellif
36	218.5	5.2	574	1	CDAS_THEET	P29964 thermoanaer
37	215.5	5.1	573	1	TRES_PIMSR	P72235 pimelobacte
38	214.5	5.1	627	1	GLGB_BACSU	P39118 bacillus su
39	212.5	5.0	581	1	MAXS_YEAST	P40884 saccharomyc
40	208	4.9	764	1	GLGB_STRAU	P52980 streptomyce
41	205.5	4.9	1279	1	APU_THESA	P36905 t amylopull
42	201	4.8	562	1	AMY2_DICTH	P14898 dictyoglomu
43	200	4.7	561	1	TREC_BACSU	P39795 bacillus su
44	200	4.7	734	1	GLGB_AGRTU	P52979 agrobacteri
45	197	4.7	1196	1	AMYB_PAEPO	P21543 paenibacill

ALIGNMENTS

```
RESULT   1
GLGX_MYCTU
ID   GLGX_MYCTU     STANDARD;      PRT;   721 AA.
AC   Q10767;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Glycogen operon protein glgX homolog (EC 3.2.1.-).
GN   GLGX OR RV1564C OR MT1615 OR MTCY48.01.
OS   Mycobacterium tuberculosis.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX   NCBI_TaxID=1773;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=H37RV;
RX   MEDLINE=98295987; PubMed=9634230;
RA   Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA   Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekaia F.,
RA   Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA   Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA   Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA   Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA   Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA   Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT   "Deciphering the biology of Mycobacterium tuberculosis from the
RT   complete genome sequence.";
RL   Nature 393:537-544(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CDC 1551 / Oshkosh;
RA   Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA   Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA   Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA   Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA   Bishai W.;
RT   "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT   laboratory strains.";
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC       KNOWN AS THE ALPHA-AMYLASE FAMILY. STRONG, TO E.COLI AND
CC       H.INFLUENZAE GLGX.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z74020; CAA98327.1; -.
DR   EMBL; AE007027; AAK45882.1; -.
DR   HSSP; P10342; 1BF2.
DR   TIGR; MT1615; -.
DR   TubercuList; Rv1564c; -.
```

```
DR   InterPro; IPR000461; Alpha_amylase.
DR   InterPro; IPR004193; isoamylase_N.
DR   Pfam; PF00128; alpha-amylase; 1.
DR   Pfam; PF02922; isoamylase_N; 1.
KW   Hydrolase; Glycosidase; Complete proteome.
FT   ACT_SITE    362    362       BY SIMILARITY.
FT   ACT_SITE    470    470       BY SIMILARITY.
SQ   SEQUENCE   721 AA;  81081 MW;  5616457CDCC887D1 CRC64;


  Query Match             29.1%;  Score 1225;  DB 1;  Length 721;
  Best Local Similarity   39.4%;  Pred. No. 1.2e-85;
  Matches  280;  Conservative   85;  Mismatches  263;  Indels   82;  Gaps   19;

Qy      57 VLAGMPAPLGATALAGGVNFAVYSGGATAAALCLFTPEDLKADRVTEEVPLDPLMNRTGN 116
           |  |   |||||    | ||:::|  |    |||     :  | |   :||| :    |
Db      19 VWPGNAYPLGATYDGAGTNFSLFSEIAEKVELCL-----IDEDGVESRIPLDEV---DGY 70

Qy     117 VWHVFIEGELHNMLYGYRFDGTFAPHCGHYLDVSNVVVDPYAKAVISRGEYG-------- 168
           ||| ::        ||:|  | | |  ||  | | :::||| |:      :|
Db      71 VWHAYLPNITPGQRYGFRVHGPFDPAAGHRCDPSKLLLDPYGKSFHGDFTFGQALYSYDV 130

Qy     169 -------VPARGNNCWPQMAGMIPLPYSTFDWEGDLPLRYPQKDLVIYEMHLRGFTK-HD 220
                 |   ::    |  ::  |:  ||| |   | |  : |||| |::| |: |
Db     131 NAVDPDSTPPMVDSLGHTMTSVVINPF--FDWAYDRSPRTPYHETVIYEAHVKGMTQTHP 188

Qy     221 SSNVEHPGTFIGAVSK--LDYLKELGVNCIELMPCHEFNELEYSTSSSKMNFWGYSTINF 278
           |   |  ||: |      :|:| || |  :|||| |:|             |:|||:|  |
Db     189 SIPPELRGTYAGLAHPVIIDHLNELNVTAVELMPVHQFLHDSRLLDLGLRNYWGYNTFGF 248

Qy     279 FSPMTRYTSGGIKNCGRDAINEFKTFVREAHKRGIEVILDVVFNHTAEGNENGPILSFKG 338
           |:|  :| |    :  |  |: |||| ||  |: |||||||||:|||||||  || ::|:|
Db     249 FAPHHQYAS--TRQAG-SAVAEFKTMVRSLHEAGIEVILDVVYNHTAEGNHLGPTINFRG 305

Qy     339 VDNTTYYMLAPKG-EFY-NYSGCGNTFNCNHPVVRQFIVDCLRYWVTEMHVDGFRFDLAS 396
           :||| || |      || :::| ||: |  ||   | |:| ||||| ||||||||||||||
Db     306 IDNTAYYRLMDHDLRFYKDFTGTGNSLNARHPHTLQLIMDSLRYWVIEMHVDGFRFDLAS 365

Qy     397 IMTRGSSLWDPVNVYGAPIEGDMITTGTPLVTPPLIDMISNDPILGGVKLIAEAWDAG-G 455
           : |     | :: :                       |::  ||::  |||||| || | |
Db     366 TLARELHDVDRLSAF-------------------FDLVQQDPVVSQVKLIAEPWDVGEG 405

Qy     456 LYQVGQFPHWNVWSEWNGKYRDIVRQFIKGTDGFAGGFAECLCGSPHLYQAGGRKPWHSI 515
           |||| ||   :|:|||||||| || : :|     | ||  | ||  ||:| ||:|  ||
Db     406 GYQVGNFP--GLWTEWNGKYRDTVRDYWRGEPATLGEFASRLTGSSDLYEATGRRPSASI 463

Qy     516 NFVCAHDGFTLADLVTYNKKYNLPNGENNRDGENHNLSWNCGEEGEFARLSVKRLRKRQM 575
           ||| ||||||| |||:|| |:|  ||||||||::| ||||| ||      :  || |||
Db     464 NFVTAHDGFTLNDLVSYNDKHNEANGENNRDGESYNRSWNCGVEGPTDDPDILALRARQM 523

Qy     576 RNFFVCLMVSQGVPMFYMGDEYGHTKGGNNNTYCHDSYVNYFRWDKKEQYSELHRFCCLM 635
           || :  |||||| ||   ||| | |: |||| || || ::::  |   :: ::|  |
Db     524 RNMWATLMVSQGTPMIAHGDEIGRTQYGNNNVYCQDSELSWMDWSLVDKNADLLAFARKA 583

Qy     636 TKFRKE---------CEGLGLEDFPTAKRLQW---HGHQPGKPDWSEN-SRFVAFSMKDE 682
           |  ||           ||  :      : : |    | :    ||     | ||  :  |
Db     584 TTLRKNHKVFRRRRFFEGEPIRSGDEVRDIAWLTPSGREMTHEDWGRGFDRCVAVFLNGE 643

Qy     683 ------RQGE------IYVAFNTSHLPAVVELP-ERAGRRWEPVVDTGKP 719
                :||        : ||         :| :   ::|   :||  |
Db     644 AITAPDARGERVVDDSFLLCFNAHDHDVEFVMPHDGYAQQWTGELDTNDP 693


RESULT   2
GLGX_HAEIN
ID   GLGX_HAEIN     STANDARD;      PRT;   659 AA.
AC   P45178;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Glycogen operon protein glgX (EC 3.2.1.-).
GN   GLGX OR HI1358.
OS   Haemophilus influenzae.
OC   Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC   Haemophilus.
OX   NCBI_TaxID=727;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=RD / KW20 / ATCC 51907;
RX   MEDLINE=95350630; PubMed=7542800;
RA   Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA   Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA   McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA   Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA   Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA   Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA   Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA   Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA   Venter J.C.;
RT   "Whole-genome random sequencing and assembly of Haemophilus
RT   influenzae Rd.";
RL   Science 269:496-512(1995).
CC   -!- FUNCTION: THIS PROTEIN MAY BE PART OF A GLYCOGEN BIOSYNTHETIC/
CC       CATABOLIC OPERON BUT IS NOT REQUIRED FOR GLYCOGEN SYNTHESIS
CC       (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC       KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U32815; AAC23005.1; -.
DR   HSSP; P10342; 1BF2.
DR   TIGR; HI1358; -.
DR   InterPro; IPR000461; Alpha_amylase.
DR   InterPro; IPR004193; isoamylase_N.
DR   Pfam; PF00128; alpha-amylase; 1.
DR   Pfam; PF02922; isoamylase_N; 1.
KW   Glycogen biosynthesis; Hydrolase; Glycosidase; Complete proteome.
FT   ACT_SITE    333    333       BY SIMILARITY.
FT   ACT_SITE    440    440       BY SIMILARITY.
SQ   SEQUENCE   659 AA;  75290 MW;  E2B71F610E8B1CAC CRC64;


  Query Match             24.1%;  Score 1014;  DB 1;  Length 659;
  Best Local Similarity   33.9%;  Pred. No. 1.4e-69;
  Matches  236;  Conservative  112;  Mismatches  242;  Indels  106;  Gaps   21;

Qy      60 GMPAPLG-ATALAGGV---NFAVYSGGATAAALCLFTPEDLKADRVTEEVPLDPLMNRTG 115
           | | |:| : |:    |   |||::|  |    ||||       |   :|  |   | ||
Db       8 GNPIPMGYSQAVENNVQITNFALFSAAAIGVELCLF-------DEQNQETRLP--MVRTE 58

Qy     116 NVWHVFIEGELHNMLYGYRFDGTFAPHCGHYLDVSNVVVDPYAKAVISRGEYGVPARGNN 175
           ||||: : |      | :|  | ||       :  :::|||||||   |:  : :  :
Db      59 NVWHLAVTGVKTGTEYAFRIHGEFA-------NPQKLILDPYAKAV--NGKPDLSSEESR 109

Qy     176 CW---------PQMAGMIPLPYSTFDWEGDLPLRYPQKDLVIYEMHLRGFTKHDSSNVEH 226
           |           :|    :    |||| |     | : ::||:|::||::    | :
Db     110 SWFLLSDNRDNAHLAPRAVVISEEFDWENDTSPNTPWAETIVYELHVKGFSQ---LNEKI 166

Qy     227 P----GTFIGAVS--KLDYLKELGVNCIELMPCHEFNELEYSTSSSKMNFWGYSTINFFS 280
           |    ||: |      | |||||||  :||:| :      :  :    |:|||: :  |:
Db     167 PAALRGTYTGLAHPVNLAYLKELGVTAVELLPVNFHINEPHLQARGLQNYWGYNPLAMFA 226

Qy     281 PMTRYTSGGIKNCGRDAINEFKTFVREAHKRGIEVILDVVFNHTAEGNENGPILSFKGVD 340
              :| :        : : |||| |:  || |||||||||||:||  :  |  | :|:|
Db     227 VEPKYAA------TNNPLAEFKTMVKAFHKAGIEVILDVVFNHSAESEQTYPTFSQRGID 280

Qy     341 NTTYYMLAPKGEFYNYSGCGNTFNCNHPVVRQFIVDCLRYWVTEMHVDGFRFDLASIMTR 400
           : |||    :| : |::|||| | :  | |:::|||||||| : |:|||||||::: |
Db     281 DQTYYWRNDQGRYINWTGCGNMLNLSSDVGRKWVVDCLRYWVEQCHIDGFRFDLATVLGR 340
```

```
Qy     401 GSSLWDPVNVYGAPIEGDMITTGTPLVTPPLIDMISNDPILGGVKLIAEAWDAGGL-YQV 459
Db     341 DTPDFNS-------------------SAQLFTDIKNEPSLQNIKLIAEPWDIGHYGYQV 380

Qy     460 GQFPHWNVWSEWNGKYRDIVRQFIKGTDGFAGGFAECLCGSPHLYQAGGRKPWHSINFVC 519
Db     381 GNFPSY--FAEWNDRFRDDLCRFWLWKSGEIGAFAERFAGSSDLFKKNDRLPHTTLNFIT 438

Qy     520 AHDGFTLADLVTYNKKYNLPNGENNRDGENHNLSWNCGEEGEFARLS------VKRLRKR 573
Db     439 AHDGFTLKDLVSYNQKHNETNGEENRDGRNENYSYNHGVEGSTESLSEPQKSAVENNRTF 498

Qy     574 QMRNFFVCLMVSQGVPMFYMGDEYGHTKGGNNNTYCHDSYVNYFRWDKKEQYSELHRFCC 633
Db     499 AQSGLLMSLLLANGTPMLLAGDEFGNTQYGNNNAYCQDNEITWLKWANFNE--ELFELTK 556

Qy     634 LMTKFRKECEGLGLEDFPTAKRLQWHGHQPGKP----DWSENSRFVAFSMKDER------ 683
Db     557 QTIALRKQIGSLNKDQWWSDENVQWL-NIVGEPMTVEDWQNQQTKALQVVLDNRWLLLIN 615

Qy     684 ---QGEIYVAFNTSHLPAVVELPERAGRRWEPVVDT 716
Db     616 AKAEGQMF------HLP---------NRKWKPQIGT 636


RESULT   3
GLGX_ECOLI
ID   GLGX_ECOLI     STANDARD;      PRT;   657 AA.
AC   P15067; P76693;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Glycogen operon protein glgX (EC 3.2.1.-).
GN   GLGX OR GLYX OR B3431.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / MG1655;
RX   MEDLINE=97426617; PubMed=9278503;
RA   Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA   Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA   Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA   Mau B., Shao Y.;
RT   "The complete genome sequence of Escherichia coli K-12.";
RL   Science 277:1453-1474(1997).
RN   [2]
RP   SEQUENCE OF 1-590 FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE=89108020; PubMed=2975249;
RA   Romeo T., Kumar A., Preiss J.;
RT   "Analysis of the Escherichia coli glycogen gene cluster suggests that
RT   catabolic enzymes are encoded among the biosynthetic genes.";
RL   Gene 70:363-376(1988).
CC   -!- FUNCTION: THIS PROTEIN MAY BE PART OF A GLYCOGEN BIOSYNTHETIC/
CC       CATABOLIC OPERON BUT IS NOT REQUIRED FOR GLYCOGEN SYNTHESIS.
CC   -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC       KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC   -!- CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A STOP-
CC       CODON IN POSITION 507.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; J01616; AAA98735.1; ALT_SEQ.
DR   EMBL; U18997; AAA58229.1; -.
DR   EMBL; AE000419; AAC76456.1; -.
DR   PIR; JT0400; BVECGX.
DR   HSSP; P10342; 1BF2.
DR   EcoGene; EG10381; glgX.
DR   InterPro; IPR000461; Alpha_amylase.
DR   InterPro; IPR004193; isoamylase_N.
DR   Pfam; PF00128; alpha-amylase; 1.
DR   Pfam; PF02922; isoamylase_N; 1.
KW   Glycogen biosynthesis; Hydrolase; Glycosidase; Complete proteome.
FT   ACT_SITE    336    336       BY SIMILARITY.
FT   ACT_SITE    443    443       BY SIMILARITY.
FT   CONFLICT    288    288       Y -> YY (IN REF. 2).
SQ   SEQUENCE   657 AA;  73576 MW;  AA9EFC1F67FD0420 CRC64;


  Query Match             23.8%;  Score 1001;  DB 1;  Length 657;
  Best Local Similarity   34.4%;  Pred. No. 1.3e-68;
  Matches  235;  Conservative   98;  Mismatches  261;  Indels   90;  Gaps   18;

Qy      60 GMPAPLGATALAGGVNFAVYSGGATAAALCLFTPEDLKADRVTEEVPLDPLMNRTGNVWH 119
Db       7 GKPAPLGAHYDGQGVNFTLFSAHAERVELCVF-------DANGQEHRYD-LPGHSGDIWH 58

Qy     120 VFIEGELHNMLYGYRFDGTFAPHCGHYLDVSNVVVDPYAKAV--------ISRGEYGVPA 171
Db      59 GYLPDARPGLRYGYRVHGPWQPAEGHRFNPAKLLIDPCARQIDGEFKDNPLLHAGHNEPD 118

Qy     172 RGNNCWPQMAGMIPLPYSTFDWEGDLPLRYPQKDLVIYEMHLRGFT-KHDSSNVEHPGTF 230
Db     119 YRDNA--AIAPKCVVVVDHYDWEDDAPPRTPWGSTIIYEAHVKGLTYLHPEIPVEIRGTY 176

Qy     231 --IGAVSKLDYLKELGVNCIELMPCHEFNELEYSTSSSKMNFWGYSTINFFSPMTRYTSG 288
Db     177 KALGHPVMINYLKQLGITALELLPVAQFASEPRLQRMGLSNYWGYNPVAMFALHPAYA-- 234

Qy     289 GIKNCGRD-AINEFKTFVREAHKRGIEVILDVVFNHTAEGNENGPILSFKGVDNTTYYML 347
Db     235 ----CSPETALDEFRDAIKALHKAGIEVILDIVLNHSAELDLDGPLFSLRGIDNRSYYWI 290

Qy     348 APKGEFYNYSGCGNTFNCNHPVVRQFIVDCLRYWVTEMHVDGFRFDLASIMTRGSSLWDP 407
Db     291 REDGDYHNWTGCGNTLNLSHPAVVDYASACLRYWVETCHVDGFRFDLAAVMGR------- 343

Qy     408 VNVYGAPIEGDMITTGTPLVTP------PLIDMISNDPILGGVKLIAEAWD-AGGLYQVG 460
Db     344 -------------------TPEFRQDAPLFTAIQNCPVLSQVKLIAEPWDIAPGGYQVG 383

Qy     461 QFPHWNVWSEWNGKYRDIVRQFIKGTDGFAGGFAECLCGSPHLYQAGGRKPWHSINFVCA 520
Db     384 NFP--PLFAEWNDHFRDAARRFWLHYDLPLGAFAGRFAASSDVFKRNGRLPSAAINLVTA 441

Qy     521 HDGFTLADLVTYNKKYNLPNGENNRDGENHNLSWNCGEEGEFARLSVKRLRKRQMRNFFV 580
Db     442 HDGFTLRDCVCFNHKHNEANGEENRDGTNNNYSNNHGKEGLGGSLDLVERRRDSIHALLT 501

Qy     581 CLMVSQGVPMFYMGDEYGHTKGGNNNTYCHDSYVNYFRWDKKEQYSELHRFCCLMTKFRK 640
Db     502 TLLLSQGTPMLLAGDEHGHSQHGNNNAYCQDNQLTWLDWSQAS--SGLTAFTAALIHLRK 559

Qy     641 ECEGLGLEDFPTAKRLQWHGHQPGKPDWSENSRFVAFSMKDERQG-----------EIYV 689
Db     560 R--------IPALVENRWWEEGDGNVRWL--NRYAQPLSTDEWQNGPKQLQILLSDRFLI 609

Qy     690 AFNTSHLPAVVELPERAGRRWEPV 713
Db     610 AINATLEVTEIVLP--AG-EWHAI 630


RESULT   4
ISOA_FLASP
ID   ISOA_FLASP     STANDARD;      PRT;   777 AA.
AC   O32611;
```

```
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   Isoamylase precursor (EC 3.2.1.68).
GN   IAM.
OS   Flavobacterium sp.
OC   Bacteria; CFB group; Flavobacteria; Flavobacteriaceae; Flavobacterium.
OX   NCBI_TaxID=239;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97340931; PubMed=9197405;
RA   Krohn B.M., Barry G.F., Kishore G.M.;
RT   "An isoamylase with neutral pH optimum from a Flavobacterium species:
RT   cloning, characterization and expression of the iam gene.";
RL   Mol. Gen. Genet. 254:469-478(1997).
CC   -!- FUNCTION: HAS A HIGH RATE OF HYDROLYSIS FOR GLYCOGEN. DOES NOT
CC       CLEAVE PULLULAN. HAS A PH OPTIMUM OF 6-7.
CC   -!- CATALYTIC ACTIVITY: Hydrolysis of alpha-(1,6)-D-glucosidic branch
CC       linkages in glycogen, amylopectin and their beta-limits dextrins.
CC   -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC       KNOWN AS THE ALPHA-AMYLASE FAMILY. ISOAMYLASE SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U90120; AAB63356.1; -.
DR   HSSP; P10342; 1BF2.
DR   InterPro; IPR000461; Alpha_amylase.
DR   InterPro; IPR004193; isoamylase_N.
DR   Pfam; PF00128; alpha-amylase; 1.
DR   Pfam; PF02922; isoamylase_N; 1.
KW   Hydrolase; Glycosidase; Signal.
FT   SIGNAL        1     32       POTENTIAL.
FT   CHAIN        33    777       ISOAMYLASE.
FT   ACT_SITE    410    410       BY SIMILARITY.
FT   ACT_SITE    478    478       BY SIMILARITY.
FT   ACT_SITE    533    533       BY SIMILARITY.
FT   DISULFID    419    423       BY SIMILARITY.
SQ   SEQUENCE   777 AA;  84340 MW;  90548988EFD1445B CRC64;

  Query Match             19.2%;  Score 810;  DB 1;  Length 777;
  Best Local Similarity   32.6%;  Pred. No. 6e-54;
  Matches  257;  Conservative   96;  Mismatches  271;  Indels  164;  Gaps   38;

Qy      42 EPVAEDRYALGGACRVL-------AGMPA-------PLGATALAGGVN--FAVYSGGATA 85
           :| | |   |   | |       |  ||        |||   |   |  | |||  ||
Db       2 DPHAPQRQRSGQRLRALALAALACALSPAHAAIDAQQLGARYDAAQANLAFRVYSSRATR 61

Qy      86 AALCLF-TPED--------LKADRVTE----EVPLDPLMNRTGNVWHVFIEGELHNMLYG 132
             : |:  |          |  |  |:     :|   : |  |      | | ::   ||
Db      62 VEVFLYKNPTGSQEVARLALSKDPATQVWSLSLPTSTIKNTYG------ITGAVY---YG 112

Qy     133 YRFDGTFAPH---------CGHYLDVSN---------VVVDPYAK--------AVISRGE 166
           ||  |   |:         |   || |         :::||||:        |  : |
Db     113 YRAWGPNWPYDAAWTKGSATGFVSDVDNAGNRFNPNKLLLDPYAREISQDPNTATCADGT 172

Qy     167 -YGVPARGNN-----CWPQMAGMIPLPYSTFDWEGDLPLRYPQKDLVIYEMHLRGFTKHD 220
            |   |   |      |  :   :      |      |  | |   || ||||:|:|| |::|
Db     173 IYATGAAHRNKDSGLCASKGIALAADATSV----GSKPTR-ALKDEVIYEVHVRGLTRND 227

Qy     221 SS-NVEHPGTFIGAVSKLDYLKELGVNCIELMPCHEF----NELEYSTSSSKMNFWGYST 275
             |      ||: ||  |   |  |||  :| :|  |     |:::   |::  |:||| |
Db     228 DSVPAAERGTYKGAARKAAALAALGVTAVEFLPVQETQNDQNDVD-PNSTAGDNYWGYMT 286

Qy     276 INFFSPMTRY----TSGGIKNCGRDAINEFKTFVREAHKRGIEVILDVVFNHTAEGNENG 331
           :|:|:|  ||    ::||          |:|  |:  |  ||:| :|||:||| ||   |
Db     287 LNYFAPDRRYAYDKSAGG-------PTREWKAMVKAFHDAGIKVYIDVVYNHTGEG---G 336

Qy     332 P----------ILSFKGVDNTTYYMLAPKGEF-YNYSGCGNTFNCNHPVVRQFIVDCLR 379
           |          :|||:|:||  || |:   :: :: :| |  :|  ||: :  ||| |
Db     337 PWSGTDGLSVYNLLSFRGLDNPAYYSLSSDYKYPWDNTGVGGNYNTRHPIAQNLIVDSLA 396

Qy     380 YWVTEMHVDGFRFDLASIMTRGSSLWDPVNVYGAPIEGDMITTGTPLVTPPLIDMISNDP 439
           ||   : |||||||||::  |:|       :     |: :           ::  :   |
Db     397 YWRDALGVDGFRFDLASVL--GNSCQHGCFNFDKNDSGNALNR--------IVAELPPRP 446

Qy     440 ILG--GVKLIAEAWDAGG-LYQVGQFPHWNVWSEWNGKYRDIVR--QFIKGTDGFA-GGF 493
             |  |  |||| |  ||  |||| ||    |:|||| ||| :|  |   | :    |
Db     447 AAGGAGADLIAEPWAIGGNSYQVGGFPAG--WAEWNGLYRDALRKKQNKLGVETVTPGTL 504

Qy     494 AECLCGSPHLYQAGGRKPWHSINFVCAHDGFTLADLVTYNKKYN---LPNGENNRDGENH 550
           |    ||  ||   |||||||||| |||||| ||  || | |    | | ::  ||:|
Db     505 ATRFAGSNDLYGDDGRKPWHSINFVVAHDGFTLNDLYAYNDKQNNQPWPYGPSD-GGEDH 563

Qy     551 NLSWNCGEEGEFARLSVKRLRKRQMRNFFVCLMVSQGVPMFYMGDEYGHTKGGNNNTYCH 610
           ||||| |         :  ::: |     ||:| ||||   |||   |: ||||||
Db     564 NLSWNQG--------GIVAEQRKAARTGLALLMLSAGVPMITGGDEALRTQFGNNNTYNL 615

Qy     611 DSYVNYFRWDKKEQYSELHRFCCLMTKFRKECEGLGLEDFPTA--------KRLQWHGHQ 662
           ||  |: | :  ::   :  : |||    |  :| :|        ::|:|
Db     616 DSAANWLYWSRSALEADHETYTKRLIAFRKAHPALRPANFYSASDTNGNVMEQLRWF--- 672

Qy     663 PGKPDWSE---------NSRFVAFSMKDERQGE----IYVAFNTSHLPAVVELP-ERAGR 708
             ||| ::          ::  :|: :     |:    ||||:|        :||    |:
Db     673 --KPDGAQADSAYFNGADNHALAWRIDGSEFGDSASAIYVAYNGWSGAVDFKLPWPGTGK 730

Qy     709 RWEPVVDT 716
           :|  | ||
Db     731 QWYRVTDT 738


RESULT   5
ISOA_PSESP
ID   ISOA_PSESP     STANDARD;.      PRT;   776 AA.
AC   P26501;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   Isoamylase precursor (EC 3.2.1.68).
GN   IAM.
OS   Pseudomonas sp. (strain SMP1).
OC   Bacteria; Proteobacteria.
OX   NCBI_TaxID=306;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 27-44.
RX   MEDLINE=89381677; PubMed=2778432;
RA   Tognoni A., Carrera P., Galli G., Lucchese G., Camerini B.,
RA   Grandi G.;
RT   "Cloning and nucleotide sequence of the isoamylase gene from a strain
RT   of Pseudomonas sp.";
RL   J. Gen. Microbiol. 135:37-45(1989).
CC   -!- CATALYTIC ACTIVITY: Hydrolysis of alpha-(1,6)-D-glucosidic branch
CC       linkages in glycogen, amylopectin and their beta-limits dextrins.
CC   -!- INDUCTION: BY MALTOSE.
CC   -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC       KNOWN AS THE ALPHA-AMYLASE FAMILY. ISOAMYLASE SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M25247; AAA25855.1; -.
DR   EMBL; A10909; CAA00929.1; -.
DR   PIR; A37035; A37035.
DR   HSSP; P10342; 1BF2.
DR   InterPro; IPR000461; Alpha_amylase.
```

```
DR   InterPro; IPR004193; isoamylase_N.
DR   Pfam; PF00128; alpha-amylase; 1.
DR   Pfam; PF02922; isoamylase_N; 1.
KW   Hydrolase; Glycosidase; Signal.
FT   SIGNAL        1     26
FT   CHAIN        27    776       ISOAMYLASE.
FT   ACT_SITE    401    401       BY SIMILARITY.
FT   ACT_SITE    481    481       BY SIMILARITY.
FT   ACT_SITE    536    536       BY SIMILARITY.
FT   DISULFID    410    422       BY SIMILARITY.
FT   DISULFID    546    616       BY SIMILARITY.
FT   DISULFID    738    766       BY SIMILARITY.
SQ   SEQUENCE   776 AA;  83656 MW;  A5E4C02EF026A3A4 CRC64;


  Query Match             19.1%;  Score 803.5;  DB 1;  Length 776;
  Best Local Similarity   32.0%;  Pred. No. 1.9e-53;
  Matches  242;  Conservative   99;  Mismatches  289;  Indels  127;  Gaps   31;

Qy      55 CRVLAGMPA----------PLGAT--ALAGGVNFAVYSGGATAAALCLFTPEDLKADRVT 102
           | ||||:||           |||:  |    : | |||  ||   | |::     |
Db      13 CAVLAGVPAMPAHAAINSMSLGASYDAQQANITFRVYSSQATRIVLYLYS-----AGYGV 67

Qy     103 EEVPLDPLMNRTGNVWHVFI-------EGELHNMLYGYRFDGTFAPHC------------ 143
           :|     |      || | :         |    : ||||  |   |:
Db      68 QESATYTLSPAGSGVWAVTVPVSSIKAAGITGAVYYGYRAWGPNWPYASNWGKGSQAGFV 127

Qy     144 ------GHYLDVSNVVVDPYAKAVISRGEYGVPARGNNCW--------------PQMAG 182
                 |   : : :::|||||: | |:        :  | :                |:
Db     128 SDVDANGDRFNPNKLLLDPYAQEV-SQDPLNPSNQNGNVFASGASYRTTDSGIYAPKGVV 186

Qy     183 MIPLPYSTFDWEGDLPLRYPQKDLVIYEMHLRGFTKHDSS-NVEHPGTFIGAVSKLDYLK 241
           ::|   ||    |  | |  ||| ||||:|:|||||: |:|   :: ||: ||  |  ||
Db     187 LVPSTQST----GTKPTR-AQKDDVIYEVHVRGFTEQDTSIPAQYRGTYYGAGLKASYLA 241

Qy     242 ELGVNCIELMPCHE----FNELEYSTSSSKMNFWGYSTINFFSPMTRY----TSGGIKNC 293
           |||  :| :|  |     |::    | :  |:||| | |:|||  ||     :||
Db     242 SLGVTAVEFLPVQETQNDANDV-VPNSDANQNYWGYMTENYFSPDRRYAYNKAAGG---- 296

Qy     294 GRDAINEFKTFVREAHKRGIEVILDVVFNHTAEGN-------ENGPILSFKGVDNTTYYM 346
           ||:  |:  |  ||:| :|||:||||||             | |::|:||||||
Db     297 ---PTAEFQAMVQAFHNAGIKVYMDVVYNHTAEGGTWTSSDPTTATIYSWRGLDNTTYYE 353

Qy     347 LAPKGE-FYNYSGCGNTFNCNHPVVRQFIVDCLRYWVTEMHVDGFRFDLASIMTRGSSLW 405
           |    : ||: :| |  ||  : | :  ||| | ||   | ||||||||||::  |:|
Db     354 LTSGNQYFYDNTGIGANFNTYNTVAQNLIVDSLAYWANTMGVDGFRFDLASVL--GNSCL 411

Qy     406 DPVNVYGAP--IEG--DMITTGTPLVTPPLIDMISNDPILG--GVKLIAEAWDAGG-LYQ 458
           :      ||    |  :     : :    ::  :  |  |  |: | || |  ||  ||
Db     412 NGAYTASAPNCPNGGYNFDAADSNVAINRILREFTVRPAAGGSGLDLFAEPWAIGGNSYQ 471

Qy     459 VGQFPHWNVWSEWNGKYRDIVRQF---IKGTDGFAGGFAECLCGSPHLYQAGGRKPWHSI 515
           :| ||    ||||||  :|| :||    :     :    |    || :|:|: || ||:||
Db     472 LGGFPQG--WSEWNGLFRDSLRQAQNELGSMTIYVTQDANDFSGSSNLFQSSGRSPWNSI 529

Qy     516 NFVCAHDGFTLADLVTYNKKYN---LPNGENNRDGENHNLSWNCGEEGEFARLSVKRLRK 572
           ||:  ||| || |: : |    |    | | ::  | : | ||   ::|  |       ::
Db     530 NFIDVHDGMTLKDVYSCNGANNSQAWPYGPSD-GGTSTNYSW---DQGMSAGTGAAVDQR 585

Qy     573 RQMRNFFVCLMVSQGVPMFYMGDEYGHTKGGNNNTYCHDSYVNYFRWDKKEQYSELHRFC 632
           |  |      :| | |:   ||||  |   ||| |  ||  |:  :      |  : |
Db     586 RAARTGMAFEMLSAGTPLMQGGDEYLRTLQCNNNAYNLDSSANWLTYSWTTDQSNFYTFA 645

Qy     633 CLMTKFRKECEGLGLEDFPTAKRLQWHGHQPGKPD---WSENSRF-VAFSMKDERQGE-- 686
             : |||    |    : :  :| |:       |   |:  | : :|:::     |:
Db     646 QRLIAFRKAHPALRPSSWYSGSQLTWYQPSGAVADSNYWNNTSNYAIAYAINGPSLGDSN 705

Qy     687 -IYVAFN------TSHLPAVVELPERAGRRWEPVVDT 716
            ||||:|      |  |||       :| :|  | ||
Db     706 SIYVAYNGWSSSSVTFTLPA-----PPSGTQWYRVTDT 737


RESULT   6
```

```
ISOA_PSEAY
ID   ISOA_PSEAY     STANDARD;      PRT;   776 AA.
AC   P10342;
DT   01-MAR-1989 (Rel. 10, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   Isoamylase precursor (EC 3.2.1.68).
GN   IAM.
OS   Pseudomonas amyloderamosa.
OC   Bacteria; Proteobacteria.
OX   NCBI_TaxID=32043;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SB-15;
RX   MEDLINE=88243808; PubMed=3379068;
RA   Amemura A., Chakraborty R., Fujita M., Noumi T., Futai M.;
RT   "Cloning and nucleotide sequence of the isoamylase gene from
RT   Pseudomonas amyloderamosa SB-15.";
RL   J. Biol. Chem. 263:9271-9275(1988).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JD210;
RX   MEDLINE=91064385; PubMed=2248978;
RA   Chen J.H., Chen Z.Y., Chow T.Y., Chen J.C., Tan S.T., Hsu W.H.;
RT   "Nucleotide sequence and expression of the isoamylase gene from an
RT   isoamylase-hyperproducing mutant, Pseudomonas amyloderamosa JD210.";
RL   Biochim. Biophys. Acta 1087:309-315(1990).
RN   [3]
RP   SEQUENCE OF 744-776 FROM N.A.
RC   STRAIN=SB-15;
RX   MEDLINE=89327147; PubMed=2753857;
RA   Amemura A., Fujita M., Futai M.;
RT   "Transcription of the isoamylase gene (iam) in Pseudomonas
RT   amyloderamosa SB-15.";
RL   J. Bacteriol. 171:4320-4325(1989).
RN   [4]
RP   X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RX   MEDLINE=98387895; PubMed=9719642;
RA   Katsuya Y., Mezaki Y., Kubota M., Matsuura Y.;
RT   "Three-dimensional structure of Pseudomonas isoamylase at 2.2-A
RT   resolution.";
RL   J. Mol. Biol. 281:885-897(1998).
CC   -!- CATALYTIC ACTIVITY: Hydrolysis of alpha-(1,6)-D-glucosidic branch
CC       linkages in glycogen, amylopectin and their beta-limits dextrins.
CC   -!- INDUCTION: BY MALTOSE.
CC   -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC       KNOWN AS THE ALPHA-AMYLASE FAMILY. ISOAMYLASE SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; J03871; AAA25854.1; -.
DR   EMBL; X13378; CAA31754.1; -.
DR   PIR; A28109; A28109.
DR   PDB; 1BF2; 12-AUG-98.
DR   InterPro; IPR000461; Alpha_amylase.
DR   InterPro; IPR004193; isoamylase_N.
DR   Pfam; PF00128; alpha-amylase; 1.
DR   Pfam; PF02922; isoamylase_N; 1.
KW   Hydrolase; Glycosidase; Signal; 3D-structure.
FT   SIGNAL        1     26
FT   CHAIN        27    776       ISOAMYLASE.
FT   ACT_SITE    401    401       BY SIMILARITY.
FT   ACT_SITE    481    481       BY SIMILARITY.
FT   ACT_SITE    536    536       BY SIMILARITY.
FT   DISULFID    410    422
FT   DISULFID    546    616
FT   DISULFID    738    766
```

```
FT   CONFLICT      8      8       A -> G (IN REF. 1).
FT   CONFLICT    126    126       F -> C (IN REF. 1).
FT   CONFLICT    169    171       GAS -> AH (IN REF. 1).
FT   CONFLICT    386    386       L -> V (IN REF. 1).
FT   CONFLICT    413    416       GAYT -> AVH (IN REF. 1).
FT   CONFLICT    454    489       SGLDLFAEPWAIGGNSYQLGGFPQGWSEWNGLFRDS ->
FT                                TVWICLRNLGPSAATRTSWVDSRRVVRVEWSVPRQ (IN
FT                                REF. 1).
FT   CONFLICT    555    556       WP -> S (IN REF. 1).
FT   CONFLICT    650    657       AFRKAHPA -> RSARHIP (IN REF. 1).
SQ   SEQUENCE   776 AA;  83626 MW;  F738BF8040246169 CRC64;


  Query Match             19.0%;  Score 798.5;  DB 1;  Length 776;
  Best Local Similarity   31.8%;  Pred. No. 4.5e-53;
  Matches  241;  Conservative   99;  Mismatches  290;  Indels  127;  Gaps   31;

Qy      55 CRVLAGMPA----------PLGAT--ALAGGVNFAVYSGGATAAALCLFTPEDLKADRVT 102
           | ||||:||           |||:  |    : | |||  ||   | |::     |
Db      13 CAVLAGVPAMPAHAAINSMSLGASYDAQQANITFRVYSSQATRIVLYLYS-----AGYGV 67

Qy     103 EEVPLDPLMNRTGNVWHVFI-------EGELHNMLYGYRFDGTFAPHC------------ 143
           :|     |      || | :        |    : ||||  |   |:
Db      68 QESATYTLSPAGSGVWAVTVPVSSIKAAGITGAVYYGYRAWGPNWPYASNWGKGSQAGFV 127

Qy     144 ------GHYLDVSNVVVDPYAKAVISRGEYGVPARGNNCW--------------PQMAG 182
                |   : : :::|||||: | |:       :  | :              |:
Db     128 SDVDANGDRFNPNKLLLDPYAQEV-SQDPLNPSNQNGNVFASGASYRTTDSGIYAPKGVV 186

Qy     183 MIPLPYSTFDWEGDLPLRYPQKDLVIYEMHLRGFTKHDSS-NVEHPGTFIGAVSKLDYLK 241
           ::|   ||     |  | |  ||| ||||:|:|||||: |:|   :: ||: ||  |  ||
Db     187 LVPSTQST----GTKPTR-AQKDDVIYEVHVRGFTEQDTSIPAQYRGTYYGAGLKASYLA 241

Qy     242 ELGVNCIELMPCHE----FNELEYSTSSSKMNFWGYSTINFFSPMTRY----TSGGIKNC 293
            |||  :| :|  |     |::    | :  |:||| | |:|||  ||     :||
Db     242 SLGVTAVEFLPVQETQNDANDV-VPNSDANQNYWGYMTENYFSPDRRYAYNKAAGG---- 296

Qy     294 GRDAINEFKTFVREAHKRGIEVILDVVFNHTAEGN-------ENGPILSFKGVDNTTYYM 346
           ||:  |:  |  ||:| :|||:||||||             | |::|:|| |||
Db     297 ---PTAEFQAMVQAFHNAGIKVYMDVVYNHTAEGGTWTSSDPTTATIYSWRGLDNATYYE 353

Qy     347 LAPKGE-FYNYSGCGNTFNCNHPVVRQFIVDCLRYWVTEMHVDGFRFDLASIMTRGSSLW 405
           |    : ||: :| |  ||   : | :  ||| | ||   | ||||||||||::  |:|
Db     354 LTSGNQYFYDNTGIGANFNTYNTVAQNLIVDSLAYWANTMGVDGFRFDLASVL--GNSCL 411

Qy     406 DPVNVYGAP--IEG--DMITTGTPLVTPPLIDMISNDPILG--GVKLIAEAWDAGG-LYQ 458
           :      ||    |  :     : :    ::   :  |  |  |: | || |  ||  ||
Db     412 NGAYTASAPNCPNGGYNFDAADSNVAINRILREFTVRPAAGGSGLDLFAEPWAIGGNSYQ 471

Qy     459 VGQFPHWNVWSEWNGKYRDIVRQF---IKGTDGFAGGFAECLCGSPHLYQAGGRKPWHSI 515
           :| ||    ||||||  :|| :||    :     :    |    || :|:|: || ||:||
Db     472 LGGFPQG--WSEWNGLFRDSLRQAQNELGSMTIYVTQDANDFSGSSNLFQSSGRSPWNSI 529

Qy     516 NFVCAHDGFTLADLVTYNKKYN---LPNGENNRDGENHNLSWNCGEEGEFARLSVKRLRK 572
           ||:  ||| || |: : |    |     | | ::  | : | ||   ::|  |       ::
Db     530 NFIDVHDGMTLKDVYSCNGANNSQAWPYGPSD-GGTSTNYSW---DQGMSAGTGAAVDQR 585

Qy     573 RQMRNFFVCLMVSQGVPMFYMGDEYGHTKGGNNNTYCHDSYVNYFRWDKKEQYSELHRFC 632
           |  |       |:| | |:   ||||  |   ||| |  ||  |:  :       |  : |
Db     586 RAARTGMAFEMLSAGTPLMQGGDEYLRTLQCNNNAYNLDSSANWLTYSWTTDQSNFYTFA 645

Qy     633 CLMTKFRKECEGLGLEDFPTAKRLQWHGHQPGKPD---WSENSRF-VAFSMKDERQGE-- 686
             :  |||    |    : :  :| |:       |   |:  | : :|:::     |:
Db     646 QRLIAFRKAHPALRPSSWYSGSQLTWYQPSGAVADSNYWNNTSNYAIAYAINGPSLGDSN 705

Qy     687 -IYVAFN------TSHLPAVVELPERAGRRWEPVVDT 716
            ||||:|      |  |||       :| :|  | ||
Db     706 SIYVAYNGWSSSVTFTLPA-----PPSGTQWYRVTDT 737


RESULT   7
PULA_THEMA
ID   PULA_THEMA     STANDARD;      PRT;   843 AA.
AC   O33840;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Pullulanase precursor (EC 3.2.1.41) (Alpha-dextrin endo-1,6-alpha-
DE   glucosidase) (Pullulan 6-glucanohydrolase).
GN   PULA OR TM1845.
OS   Thermotoga maritima.
OC   Bacteria; Thermotogales; Thermotoga.
OX   NCBI_TaxID=2336;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MSB8 / DSM 3109;
RX   MEDLINE=98115241; PubMed=9453151;
RA   Bibel M., Brettl C., Gosslar U., Kriegshaeuser G., Liebl W.;
RT   "Isolation and analysis of genes for amylolytic enzymes of the
RT   hyperthermophilic bacterium Thermotoga maritima.";
RL   FEMS Microbiol. Lett. 158:9-15(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MSB8 / DSM 3109;
RX   MEDLINE=99287316; PubMed=10360571;
RA   Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA   Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA   McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA   Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA   Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA   Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT   "Evidence for lateral gene transfer between Archaea and Bacteria from
RT   genome sequence of Thermotoga maritima.";
RL   Nature 399:323-329(1999).
CC   -!- CATALYTIC ACTIVITY: Hydrolysis of (1->6)-alpha-D-glucosidic
CC       linkages in pullulan and in amylopectin and glycogen, and the
CC       alpha-and beta-limit dextrins of amylopectin and glycogen.
CC   -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC       KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AJ001087; CAA04522.1; -.
DR   EMBL; AE001821; AAD36907.1; -.
DR   TIGR; TM1845; -.
DR   InterPro; IPR000461; Alpha_amylase.
DR   InterPro; IPR004193; isoamylase_N.
DR   Pfam; PF00128; alpha-amylase; 1.
DR   Pfam; PF02922; isoamylase_N; 1.
KW   Hydrolase; Glycosidase; Signal; Complete proteome.
FT   SIGNAL        1     19       POTENTIAL.
FT   CHAIN        20    843       PULLULANASE.
FT   ACT_SITE    535    535       BY SIMILARITY.
FT   ACT_SITE    564    564       BY SIMILARITY.
FT   ACT_SITE    652    652       BY SIMILARITY.
SQ   SEQUENCE   843 AA;  96261 MW;  C42DDE233D54FE77 CRC64;


  Query Match             12.4%;  Score 523;  DB 1;  Length 843;
  Best Local Similarity   26.0%;  Pred. No. 5.1e-32;
  Matches  197;  Conservative  107;  Mismatches  264;  Indels  190;  Gaps   35;

Qy      27 VVEAATKVEDEGEEDEPVAE------------DRYALGGACRVLAGMPAPLGATALAGG 73
           |:  : | ||  :: | : |             | |   |           |||
Db     184 VLSESLKEEDLRKDVELIIEGYKPARVIMMEILDDYYYDG----------ELGAVYSPEK 233

Qy      74 VNFAVYSGGATAAALCLF-TPEDLKADRVTEEVPLDPLMNRTGN-VWHVFIEGELHNMLY 131
             | |:|  :    : ||   || :  :|         |   || ||   :||:|  : |
Db     234 TIFRVWSPVSKWVKVLLFKNGEDTEPYQVVN------MEYKGNGVWEAVVEGDLDGVFY 286
```

```
Qy    132 GYRFDGTFAPHCGHYLDVSNVVVDPYAKAVISRGEYGVP---ARGNNCWPQMAGMIPLPY 188
Db    287 LYQLE--------NYGKI-RTTVDPYSKAVYANSKKSAVVNLARTN---PE--------- 325

Qy    189 STFDWEGDLPLRYPQ----KDLVIYEMHLRGFTKHDSSNVEHPGTFIGAVSK-------- 236
Db    326 ---GWEND---RGPKIEGYEDAIIYEIHIADITGLENSGVKNKGLYLGLTEENTKGPGGV 379

Qy    237 ---LDYLKELGVNCIELMPCHEF---NELEYSTSSSKMNFWGYSTINFFSPMTRYTSGGI 290
Db    380 TTGLSHLVELGVTHVHILPFFDFYTGDELD--KDFEKYYNWGYDPYLFMVPEGRYSTDP- 436

Qy    291 KNCGRDAINEFKTFVREAHKRGIEVILDVVFNHTAEGNENGPILSFKGVDNTT---YYML 347
Db    437 KN-PHTRIREVKEMVKALHKHGIGVIMDMVFPHTYGIGE------LSAFDQTVPYYFYRI 489

Qy    348 APKGEFYNYSGCGNTFNCNHPVVRQFIVDCLRYWVTEMHVDGFRFDLASIMTRGSSL--- 404
Db    490 DKTGAYLNESGCGNVIASERPMMRKFIVDTVTYWVKEYHIDGFRFDQMGLIDKKTMLEVE 549

Qy    405 -----WDP-VNVYGAPIEGDMITTGTPLVTPPLIDMISNDPILGGVKLIAEAWDAGGLYQ 458
Db    550 RALHKIDPTIILYGEPWGG-----------------WGAPIRFGKS------DVAGTHV 585

Qy    459 VGQFPHWNVWSEWNGKYRDIVR--QFIKGTDGFA-GGFAE------CLCGSPH----LYQ 505
Db    586 ----------AAFNDEFRDAIRGSVFNPSVKGFVMGGYGKETKIKRGVVGSINYDGKLIK 635

Qy    506 AGGRKPWHSINFVCAHDGFTLADLVTYNKKYNLPNGENNRDGENHNLSWNCGEEGEFARL 565
Db    636 SFALDPEETINYAACHDNHTL-----WDKNYLAAKADKKKEWTEEEL------------- 677

Qy    566 SVKRLRKRQMRNFFVCLMVSQGVPMFYMGDEYGHTKGGNNNTYCHDSYVNYFRWDKKEQY 625
Db    678 ------KNAQKLAGAILLTSQGVPFLHGGQDFCRTKNFNDNSYNAPISINGFDYERKLQF 731

Qy    626 SELHRFCCLMTKFRKECEGLGLEDFPTAKRLQWH-GHQPGKPDWSENSRFVAFSMKDERQ 684
Db    732 IDVFNYHKGLIKLRKEHPAFRLKN---AEEIKKHLEFLPG------GRRIVAFMLKDHAG 782

Qy    685 G----EIYVAFNTSHLPAVVELPERAGRRWEPVVDTGK 718
Db    783 GDPWKDIVVIYNGNLEKTTYKLPE---GKWNVVVNSQK 817
```

```
RESULT   8
O16G_BACTR
ID   O16G_BACTR     STANDARD;      PRT;   562 AA.
AC   P29094;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Oligo-1,6-glucosidase (EC 3.2.1.10) (Oligosaccharide alpha-1,6-
DE   glucosidase) (Sucrase-isomaltase) (Isomaltase) (Dextrin 6-alpha-D-
DE   glucanohydrolase).
GN   MALL.
OS   Bacillus thermoglucosidasius.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Bacillus.
OX   NCBI_TaxID=1426;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=KP1006;
RX   MEDLINE=92105078; PubMed=1761534;
RA   Watanabe K., Chishiro K., Kitamura K., Suzuki Y.;
RT   "Proline residues responsible for thermostability occur with high
RT   frequency in the loop regions of an extremely thermostable oligo-1,6-
RT   glucosidase from Bacillus thermoglucosidasius KP1006.";
RL   J. Biol. Chem. 266:24287-24294(1991).
CC   -!- CATALYTIC ACTIVITY: Hydrolysis of 1,6-alpha-D-glucosidic linkages
CC       in some oligosaccharides produced from starch and glycogen by
CC       alpha-amylase, and in isomaltose.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- MISCELLANEOUS: THIS ENZYME IS EXTREMELY THERMOSTABLE.
CC   -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC       KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D10487; BAA01368.1; -.
DR   PIR; A41707; A41707.
DR   HSSP; P21332; 1UOK.
DR   InterPro; IPR000461; Alpha_amylase.
DR   Pfam; PF00128; alpha-amylase; 1.
KW   Hydrolase; Glycosidase.
FT   ACT_SITE    199    199       BY SIMILARITY.
FT   ACT_SITE    256    256       BY SIMILARITY.
FT   ACT_SITE    330    330       BY SIMILARITY.
SQ   SEQUENCE   562 AA;  66505 MW;  925EB5924726D42A CRC64;
```

```
  Query Match             7.4%;  Score 310.5;  DB 1;  Length 562;
  Best Local Similarity   21.6%;  Pred. No. 4.7e-16;
  Matches  144;  Conservative   98;  Mismatches  168;  Indels  257;  Gaps   32;

Qy    204 KDLVIYEMHLRGFTKHDSSNVEHPGTFIGAVSKLDYLKELGVNCIELMPCHEFNELEYST 263
Db      7 KEAVVYQIYPRSFY---DSNGDGIGDIRGIIAKLDYLKELGVDVVWLSPVYK-------- 55

Qy    264 SSSKMNFWGYSTINFFSPMTRYTSGGIKNCGRDAINEF------KTFVREAHKRGIEVIL 317
Db     56 SPNDDN--GYDISDY---------------RDIMDEFGTMADWKTMLEEMHKRGIKLVM 97

Qy    318 DVVFNHTAEGNENGP--ILSFKGVDN--TTYYMLAP---------------------- 349
Db     98 DLVVNHTSDEH---PWFIESRKSKDNPYRDYYIWRPGKNGKEPNNWESVFSGSAWEYDEM 154

Qy    350 KGEFY--NYSGCGNTFNCNHPVVRQFIVDCLRYWVTEMHVDGFRFDLASIMTRGSSLWDP 407
Db    155 TGEYYLHLFSKKQPDLNWENPKVRREVYEMMKFWL-DKGVDGFRMDVINMISKVPELPD- 212

Qy    408 VNVYGAPIEG------------------------------DMITTG-TPLVTPPLIDMI 435
Db    213 ----GEPQSGKKYASGSRYYMNGPRVHEFLQEMNREVLSKYDIMTVGETPGVTPK-EGIL 267

Qy    436 SNDPILGGVKLIAEAWDAGGLYQVGQFPHWNVWSEWNGKYRDIVRQFIKGTDGFAGGFAE 495
Db    268 YTDP------------SRRELNMVFQFEHMDLDSGPGGKW-DI---------------- 297

Qy    496 CLCGSPHLYQAGGRKPWHSINFVCAHDGFTLADLVTYNKKYNLPNGENNRDGENHNLSWN 555
Db    298 --------------RPW-----------SLADL-----KKTMTKWQKELEGK----GWN 322

Qy    556 C--------------GEEGEFARLSVKRLRKRQMRNFFVCLMVSQGVPMFYMGDEYGHT 600
Db    323 SLYLNNHDQPRAVSRFGDDGKY------RVESAKMLATPLHMM--QGTPYIYQGEEIGMT 374

Qy    601 KGGNNNTYCHDSYVNYFRWDKKEQYSELHRFCCLMTKFRKECEGLGLEDFPTAKRLQWHG 660
Db    375 N--------------VRFPSIEDYRDIE----TLNMYKERVEEYGEDPQEVMEKIYYKG 415

Qy    661 HQ----PGKPDWSENSRFVAFSMKDERQGEIYVAFNTSH----LPAVVELPERAGRRWEP 712
Db    416 RDNARTPMQWDDSENAGFTA--------GTPWIPVNPNYKEINVKAALEDPNSVFHYYKK 467

Qy    713 VVDTGKP------APYDFLTDDLP-----------DRALTIHQFS----------HFLYS 745
Db    468 LIQLRKQHDIIVYGTYDLILEDDPYIYRYTRTLGNEQLIVITNFSEKTPVFRLPDHIIYK 527

Qy    746 NLYPMLS 752
```

```
            ::|
Db    528 TKELLIS 534


RESULT   9
TREZ_ARTSP
ID   TREZ_ARTSP     STANDARD;      PRT;   598 AA.
AC   Q44316;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Malto-oligosyltrehalose trehalohydrolase (EC 3.2.1.141) (MTHase) (4-
DE   alpha-D-{(1->4)-alpha-D-glucano}trehalose trehalohydrolase)
DE   (Maltooligosyl trehalose trehalohydrolase).
GN   TREZ.
OS   Arthrobacter sp. (strain Q36).
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Micrococcineae; Micrococcaceae; Arthrobacter.
OX   NCBI_TaxID=1667;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96195835; PubMed=8605217;
RA   Maruta K., Hattori K., Nakada T., Kubota M., Sugimoto T., Kurimoto M.;
RT   "Cloning and sequencing of trehalose biosynthesis genes from
RT   Arthrobacter sp. Q36.";
RL   Biochim. Biophys. Acta 1289:10-13(1996).
CC   -!- CATALYTIC ACTIVITY: Hydrolysis of alpha-(1->4)-D-glucosidic
CC       linkage in 4-alpha-D-{(1->4)-alpha-D-glucanosyl}n trehalose to
CC       yield trehalose and alpha-(1->4)-D-glucan.
CC   -!- PATHWAY: Trehalose biosynthesis.
CC   -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC       KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D63343; BAA09668.1; -.
DR   InterPro; IPR000461; Alpha_amylase.
DR   InterPro; IPR004193; isoamylase_N.
DR   Pfam; PF00128; alpha-amylase; 1.
DR   Pfam; PF02922; isoamylase_N; 1.
KW   Hydrolase; Glycosidase.
FT   ACT_SITE    267    267       BY SIMILARITY.
SQ   SEQUENCE   598 AA;  65831 MW;  8B5C610AD3766947 CRC64;


  Query Match             7.4%;  Score 310;  DB 1;  Length 598;
  Best Local Similarity   23.9%;  Pred. No. 5.6e-16;
  Matches  169;  Conservative   63;  Mismatches  262;  Indels  212;  Gaps   31;

Qy      53 GACRVLAGMPAPLGATALAGGVNFAVYSGGATAAALCLFTPEDLKADRVTEEVPLDPLMN 112
           |  |     |     | ||||  :|:     |        |||
Db      15 GPARYDVWAPNAESVTLLAGGERYAMQRRAETG-------PED---------------- 50

Qy     113 RTGNVWHVFIEGELHNMLYGYRFDGTFAPHCGHYLDVSNVVVDPYAKAVISRGEYGVPAR 172
               |         |: |||  ||   |           : ||  :    |   || |
Db      51 --AGWWTAAGAPTDGNVDYGYLLDGDETP----------LPDPRTR----RQPDGVHAL 93

Qy     173 GNNCWPQMAGMIPLPYSTFDWEGDLPLRYPQKDLVIYEMHLRGFTKHDSSNVEHPGTFIG 232
               |          | : |: |       :  ||||:||  ||         ||
Db      94 SRTFDP----------SAYSWQDDAWQGRELQGAVIYELHLGTFTPE--------GTLEA 135

Qy     233 AVSKLDYLKELGVNCIELMPCHEFNELEYSTSSSKMNFWGYSTINFFSPMTRYTSGGIKN 292
           |  |||||  |||: |||:| : ||          : |||  : :|:    |  || :
Db     136 AAGKLDYLAGLGVDFIELLPVNAFN---------GTHNWGYDGVQWFAVHEAY--GGPE- 183

Qy     293 CGRDAINEFKTFVREAHKRGIEVILDVVFNHTA-EGN---ENGPILSFKGVDNTTYYMLA 348
                  :: ||  ||  |: || |||:||    ||     || |  :|  ||

Db     184 -------AYQRFVDAAHAAGLGVIQDVVYNHLGPSGNYLPRFGPYLK-QGEGNTW----- 230

Qy     349 PKGEFYNYSGCGNTFNCNHPVVRQFIVDCLRYWVTEMHVDGFRFDLASIMTRGSSLWDPV 408
             |:  |  | |:  .  :|  ||::|:| |  |: :  |||  | |    : :       :
Db     231 --GDSVNLDGPGS----DH--VRRYILDNLAMWLRDYRVDGLRLDAVHAL-KDERAVHIL 281

Qy     409 NVYGAPIEGDMITTGTPLVTPPLIDMISNDPILGGVKLIAEAWDAGGLYQVGQFP----- 463
             :||  :       | ||       |:  |:| |          |  |    ||:
Db     282 EDFGALADQISAEVGRPLTLIAESDL--NNPRL------LYPRDVNGYGLEGQWSDDFHH 333

Qy     464 --HWNVWSEWNGKYRD------IVRQFIKG--TDGFAGGFAECLCGSPHLYQAGGRKPWH 513
             | ||  |  | | |       : :    |   ||    | |   | |
Db     334 AVHVNVTGETTGYYSDFDSLAALAKVLRDGFFHDGSYSSFRERHHGRP------------ 381

Qy     514 SINFVCAHDGFTLADLVTYNKKYNLPNGENNRDGENHNLSWNCGEEGEFARLSVKRLRKR 573
           |||   |     | ||  :               :||:     |      ||| : |
Db     382 -INFSAVHP----AALVVCS-------------QNHD---QIGNRATGDRLS-QTLPYG 418

Qy     574 QMRNFFVCLMVSQGVPMFYMGDEYG----------HTKGGNNNTYCHDSYVNYFR--WD- 620
             :    |  :      ||  ||:|||          | :              : |  ||
Db     419 SLALAAVLTLTGPFTPMLLMGEEYGASTPWQFFTSHPEPELGKATAEGRIKEFERMGWDP 478

Qy     621 ------------------------KKEQYSELHRFCCLMTKFRKECEGLGLEDFPTAKR 655
                                    :  ||:|    : :   :   || ||   |
Db     479 AVVPDPQDPETFRRSKLDWAEAAEGDHARLLELYRSLTALRRSTPDLTKLGFEDTQVA-- 536

Qy     656 LQWHGHQPGKPDWSENSRFVAFSMKDERQGEIYVAFNTSHLPAVVE 701
                   : |::|:: |     |:| : |  | |  |  ::
Db     537 -----------FDEDARWLRF-----RRGGVQVLLNFSEQPVSLD 565


RESULT  10
TREZ_RHISP
ID   TREZ_RHISP     STANDARD;      PRT;   596 AA.
AC   Q53238;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Malto-oligosyltrehalose trehalohydrolase (EC 3.2.1.141) (MTHase) (4-
DE   alpha-D-{(1->4)-alpha-D-glucano}trehalose trehalohydrolase)
DE   (Maltooligosyl trehalose trehalohydrolase).
GN   TREZ.
OS   Rhizobium sp. (strain M-11).
OC   Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC   Rhizobiaceae; Rhizobium.
OX   NCBI_TaxID=391;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96219094; PubMed=8829547;
RA   Maruta K., Hattori K., Nakada T., Kubota M., Sugimoto T., Kurimoto M.;
RT   "Cloning and sequencing of trehalose biosynthesis genes from Rhizobium
RT   sp. M-11.";
RL   Biosci. Biotechnol. Biochem. 60:717-720(1996).
CC   -!- CATALYTIC ACTIVITY: Hydrolysis of alpha-(1->4)-D-glucosidic
CC       linkage in 4-alpha-D-{(1->4)-alpha-D-glucanosyl}(n) trehalose to
CC       yield trehalose and alpha-(1->4)-D-glucan.
CC   -!- PATHWAY: Trehalose biosynthesis.
CC   -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC       KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D78001; BAA11187.1; -.
DR   InterPro; IPR000461; Alpha_amylase.
DR   Pfam; PF00128; alpha-amylase; 1.
KW   Hydrolase; Glycosidase.
FT   ACT_SITE    265    265       BY SIMILARITY.
```

```
SQ   SEQUENCE   596 AA;  65262 MW;  73EA80AE0534DDCD CRC64;

  Query Match             6.7%;  Score 283.5;  DB 1;  Length 596;
  Best Local Similarity   27.0%;  Pred. No. 5.8e-14;
  Matches  138;  Conservative   53;  Mismatches  147;  Indels  173;  Gaps   31;

Qy    131 YGYRFDGTFAPHCGHYLDVSNVVVDPYAKAVISRGEYGVPA------RGNNCWPQMAGMI 184
Db     65 YGYLLDGDEIP-----------LPDPRTR----RQPEGVHALSRTFDPGAHRW-QDAG-- 106

Qy    185 PLPYSTFDWEGDLPLRYPQKDL---VIYEMHLRGFTKHDSSNVEHPGTFIGAVSKLDYLK 241
Db    107 --------WQG--------RELQGSVIYELHIGTFTPE--------GTLDAAAGKLDYLA 142

Qy    242 ELGVNCIELMPCHEFNELEYSTSSSKMNFWGYSTINFFSPMTRYTSGGIKNCGRDAINEF 301
Db    143 GLGIDFIELLPVNAFN---------GTHNWGYDGVQWFAVHEGY--GGPA--------AY 183

Qy    302 KTFVREAHKRGIEVILDVVFNHTA-EGN---ENGPILSFKGVDNTTYYMLAPKGEFYNYS 357
Db    184 QRFVDAAHAAGLGVIQDVVYNHLGPSGNYLPRYGPYLK-HGEGNTW-------GDSVNLD 235

Qy    358 GCGNTFNCNHPVVRQFIVDCLRYWVTEMHVDGFRFDLASIMTRGSSLWDPVNVYGAPIEG 417
Db    236 GPGS----DH--VRQYILDNVAMWLRDYRVDGLRLDAVHAL-KDERAVHILEEFGA---- 284

Qy    418 DMITTGTPLVTPPLIDMISNDPILGG--VKLIAEA----------WDAGGLYQVGQFP-- 463
Db    285 ------------LADALSSE---GGRPLTLIAESDLNNPRLLYPRDVNGYGLAGQWSDD 328

Qy    464 -----HWNVWSEWNGKYRD---------IVRQFIKGTDGFAGGFAECLCGSPHLYQAGGR 509
Db    329 FHHAVHVNVSGETTGYYSDFDSLGALAKVLR------DGF---FHD---GSYSSFR--GR 374

Qy    510 KPWHSINFVCAHDGFTLADLVTYNKKYNLPNGENNRDGENHNLSWNCGEEGEFARLSVKR 569
Db    375 CHGRPINFSAVHP----AALVVCSQNHDQIGNRATGDRLSQSLPY-----GSLALAAVLT 425

Qy    570 LRKRQMRNFFVCLMVSQGVPMFYMGDEYGHT 600
Db    426 L------------TGPFTPMLFMGEEYGAT 443


RESULT  11
TREZ_ARTRM
ID   TREZ_ARTRM     STANDARD;      PRT;   575 AA.
AC   Q9AJN6;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Malto-oligosyltrehalose trehalohydrolase (EC 3.2.1.141) (MTHase) (4-
DE   alpha-D-{(1->4)-alpha-D-glucano}trehalose trehalohydrolase)
DE   (Maltooligosyl trehalose trehalohydrolase).
GN   TREZ.
OS   Arthrobacter ramosus.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Micrococcineae; Micrococcaceae; Arthrobacter.
OX   NCBI_TaxID=1672;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S34;
RA   Yamamoto T., Maruta K., Watanabe H., Yamashita H., Kubota M.,
RA   Fukuda S., Kurimoto M.;
RT   "Trehalose producing operon treYZ from Arthrobacter ramosus S34.";
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- CATALYTIC ACTIVITY: Hydrolysis of alpha-(1->4)-D-glucosidic
CC       linkage in 4-alpha-D-[(1->4)-alpha-D-glucanosyl](n) trehalose to
CC       yield trehalose and alpha-(1->4)-D-glucan.
CC   -!- PATHWAY: Trehalose biosynthesis.
CC   -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC       KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AB045141; BAB40766.1; -.
DR   InterPro; IPR000461; Alpha_amylase.
DR   Pfam; PF00128; alpha-amylase; 1.
KW   Hydrolase; Glycosidase.
FT   ACT_SITE    250    250       BY SIMILARITY.
SQ   SEQUENCE   575 AA;  63079 MW;  284A3F20207E228B CRC64;

  Query Match             6.6%;  Score 278;  DB 1;  Length 575;
  Best Local Similarity   25.9%;  Pred. No. 1.5e-13;
  Matches  136;  Conservative   56;  Mismatches  178;  Indels  156;  Gaps   29;

Qy    104 EVPLDPLMNRTGNVWHVF---IEGELHNMLYGYRFDGTFAPHCGHYLDVSNVVVDPYAKA 160
Db     25 ELPL----TRDENGWWALQQPWDGGPDLVDYGYLVDGK---------------GPFADP 64

Qy    161 VISRGEYGVPARGNNCWPQMAGMIPLPYSTFD--WEGDLPLRYPQKDL---VIYEMHLRG 215
Db     65 RSLRQPRGVHELGRE-------FDPARYAWGDDGWRG--------RDLTGAVIYELHVGT 109

Qy    216 FTKHDSSNVEHPGTFIGAVSKLDYLKELGVNCIELMPCHEFNELEYSTSSSKMNFWGYST 275
Db    110 FTPE--------GTLDSAIRRLDHLVRLGVDAVELLPVNAFN---------GTHGWGYDG 152

Qy    276 INFFSPMTRYTSGGIKNCGRDAINEFKTFVREAHKRGIEVILDVVFNHTA-EGN---ENG 331
Db    153 VLWYAVHEPY--GGPE--------AYQRFVDACHARGLAVVQDVVYNHLGPSGNHLPDFG 202

Qy    332 PILSFKGVDNTTYYMLAPKGEFYNYSGCGNTFNCNHPV---VRQFIVDCLRYWVTEMHVD 388
Db    203 PYLG-SGAANT---------------WGDALNLDGPLSDEVRRYIIDNAVYWLRDMHAD 245

Qy    389 GFRFDLASIM--TRGSSLWDPVNVYGAPIEGDMITTGTPLVTPPLIDMISNDPIL----- 441
Db    246 GLRLDAVHALRDARALHLLEELAARVDELAGEL---GRPLTLIAESDL--NDPKLIRSRA 300

Qy    442 -GGVKLIAEAWDAGGLYQVGQFPHWNVWSEWNGKYRD------IVRQFIKG--TDGFAGG 492
Db    301 AHGYGLDAQ-WDD----DVHHAVHANVTGETVGYYADFGGLGALVKVFQRGWFHDGTWSS 355

Qy    493 FAECLCGSPHLYQAGGRKPWHSINFVCAHDGFTLADLVTYNKKYNLPNGENNRDGENHNL 552
Db    356 FRERHHGRPLDPDIPFRR---LVAFAQDHDQV-----------------GNR-AVGDRM 393

Qy    553 SWNCGEEGEFARLSVKRLRKRQMRNFFVCLMVSQGVPMFYMGDEYG 598
Db    394 SAQVG-EGSLAAAA------------ALVLLGPFTPMLFMGEEWG 425


RESULT  12
TREZ_BREHE
ID   TREZ_BREHE     STANDARD;      PRT;   589 AA.
AC   O52520;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Malto-oligosyltrehalose trehalohydrolase (EC 3.2.1.141) (MTHase) (4-
DE   alpha-D-{(1->4)-alpha-D-glucano}trehalose trehalohydrolase)
DE   (Maltooligosyl trehalose trehalohydrolase).
GN   TREZ.
OS   Brevibacterium helvolum.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Micrococcineae; Brevibacteriaceae; Brevibacterium.
OX   NCBI_TaxID=1704;
RN   [1]
```

```
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 11822;
RA   Kwon T., Kim C.H., Choi Y.D.;
RT   "Maltooligosyl trehalose synthase/trehalohydrolase encode proteins for
RT   trehalose production in Brevibacterium helvolum.";
RL   Submitted (DEC-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- CATALYTIC ACTIVITY: Hydrolysis of alpha-(1->4)-D-glucosidic
CC       linkage in 4-alpha-D-[(1->4)-alpha-D-glucanosyl](n) trehalose to
CC       yield trehalose and alpha-(1->4)-D-glucan.
CC   -!- PATHWAY: Trehalose biosynthesis.
CC   -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC       KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF039919; AAB95369.1; -.
DR   InterPro; IPR000461; Alpha_amylase.
DR   Pfam; PF00128; alpha-amylase; 1.
KW   Hydrolase; Glycosidase.
FT   ACT_SITE    258    258       BY SIMILARITY.
SQ   SEQUENCE   589 AA;  64217 MW;  C282314B6E9BA029 CRC64;


  Query Match             6.6%;  Score 277.5;  DB 1;  Length 589;
  Best Local Similarity   26.9%;  Pred. No. 1.6e-13;
  Matches  132;  Conservative   61;  Mismatches  161;  Indels  137;  Gaps   29;

Qy    131 YGYRFDGTFAPHCGHYLDVSNVVVDPYAKAVISRGEYGVPARGNNCWPQMAGMIPLPYST 190
          |||  ||   |           | :| ::         :||  :|          | ||
Db     58 YGYLLDGNTTP----------VPEPRSRR--------LPAGVHN---HSRTYNPPPYRW 95

Qy    191 FD--WEGDLPLRYPQKDL---VIYEMHLRGFTKHDSSNVEHPGTFIGAVSKLDYLKELGV 245
          |  | |        |:|   :||::|: | :  |       ||   |  || || :||:
Db     96 QDSRWRG--------KELQGTLIYQLHV-GTSTPD-------GTLDAAGEKLSYLVDLGI 139

Qy    246 NCIELMPCHEFNELEYSTSSSKMNFWGYSTINFFSPMTRYTSGGIKNCGRDAINEFKTFV 305
          : |||:| : ||          : |||  : :::    |  ||          :: ||
Db    140 DFIELLPVNGFN---------GTHNWGYDGVQWYTVHEGY--GGPA--------AYQRFV 180

Qy    306 REAHKRGIEVILDVVFNHTA-EGN---ENGPILSFKGVDNTTYYMLAPKGEFYNYSGCGN 361
            ||  |: || |||:||    ||   : || |  :|  ||        |:  |  | |:
Db    181 DAAHAAGLGVIQDVVYNHLGLRGNYFPKLGPNLK-QGDANTL-------GDSVNLDGAGS 232

Qy    362 TFNCNHPVVRQFIVDCLRYWVTEMHVDGFRFDLASIM--TRGSSLWDPVNVYGAPIEGDM 419
              | |::|:|    || : ||||  ||    :  |   : : :   |  | |:
Db    233 ------DVFREYILDNAALWVGDYHVDGVGFDAVHAVRDERAVHILEDLGALGDAISGE- 285

Qy    420 ITTGTPLVTPPLIDMISNDPILGGVKLIAEAWDAGGLYQVGQFP-------HWNVWSEWN 472
            || |     : :   |:|     :||    |  |    ||:        | :|  |
Db    286 --TGLPKTL--IAESDFNNP-----RLIYPR-DVNGYGLAGQWSDDFHTAVHVSVSGETT 335

Qy    473 GKYRD-----IVRQFIKGTDGFAGGFAECLCGSPHLYQAGGRKPWHSINFVCAHDGFTLA 527
          | | |     :: : :|  |||          ||   ::  ||     ||   |:     |
Db    336 GYYSDFESLAVLAKVLK--DGFLHD------GSYSSFR--GRHHGRPINPSLANP----A 381

Qy    528 DLVTYNKKYNLPNGENNRDGENHNLSWNCGEEGEFARLSVKRLRKRQMRNFFVCLMVSQG 587
            ||  |: ::        |  : :||:     |: |  :|  |                |
Db    382 ALVVCNQNHDQIGNRATGDRLSQSLSY-----GQLAVAAVLTL------------TSPF 423

Qy    588 VPMFYMGDEYG 598
            || :||:|||
Db    424 TPMLFMGEEYG 434


RESULT  13
TREZ_MYCTU
ID   TREZ_MYCTU     STANDARD;      PRT;   580 AA.
AC   Q10769;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Malto-oligosyltrehalose trehalohydrolase (EC 3.2.1.141) (MTHase) (4-
DE   alpha-D-{(1->4)-alpha-D-glucano}trehalose trehalohydrolase)
DE   (Maltooligosyl trehalose trehalohydrolase).
GN   TREZ OR RV1562C OR MT1613 OR MTCY48.03.
OS   Mycobacterium tuberculosis.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX   NCBI_TaxID=1773;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=H37RV;
RX   MEDLINE=98295987; PubMed=9634230;
RA   Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA   Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekaia F.,
RA   Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA   Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA   Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA   Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA   Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA   Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT   "Deciphering the biology of Mycobacterium tuberculosis from the
RT   complete genome sequence.";
RL   Nature 393:537-544(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CDC 1551 / Oshkosh;
RA   Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA   Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA   Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA   Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA   Bishai W.;
RT   "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT   laboratory strains.";
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- CATALYTIC ACTIVITY: Hydrolysis of alpha-(1->4)-D-glucosidic
CC       linkage in 4-alpha-D-[(1->4)-alpha-D-glucanosyl](n) trehalose to
CC       yield trehalose and alpha-(1->4)-D-glucan.
CC   -!- PATHWAY: Trehalose biosynthesis.
CC   -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC       KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z74020; CAA98329.1; -.
DR   EMBL; AE007027; AAK45880.1; ALT_INIT.
DR   TIGR; MT1613; -.
DR   TubercuList; Rv1562c; -.
DR   InterPro; IPR000461; Alpha_amylase.
DR   InterPro; IPR004193; isoamylase_N.
DR   Pfam; PF00128; alpha-amylase; 1.
DR   Pfam; PF02922; isoamylase_N; 1.
KW   Hydrolase; Glycosidase; Complete proteome.
FT   ACT_SITE    247    247       BY SIMILARITY.
SQ   SEQUENCE   580 AA;  64076 MW;  506190468F44B862 CRC64;


  Query Match             6.5%;  Score 272;  DB 1;  Length 580;
  Best Local Similarity   21.3%;  Pred. No. 4.2e-13;
  Matches  144;  Conservative   83;  Mismatches  220;  Indels  230;  Gaps   36;

Qy    118 WHVFIEGELHNMLYGYRFDGTFAPHCGHYLDVSNVVVDPYAKAVISRGEYGVPARGNNCW 177
          ||  :     : |||  |           |   |: || :    :|   || ||
Db     33 WHTTVAAPA-DARYGYLLD----------DDPTVLPDPRS----ARQPDGVHARSQRWE 76
```

```
Qy    178 PQMAGMIPLPYSTFDWEGDLPLRYPQKDLVIYEMHLRGFTKHDSSNVEHPGTFIGAVSKL 237
Db     77 P--PGQFGAARTDTGWPG-----RSVEGAVIYELHIGTFTT--------AGTFDAAIEKL 121

Qy    238 DYLKELGVNCIELMPCHEFNELEYSTSSSKMNFWGYSTINFFSPMTRYTSGGIKNCGRDA 297
Db    122 DYLVDLGIDFVELMPVNSF---------AGTRGWGYDGVLWYSVHEPY--GGPDGLVR-- 168

Qy    298 INEFKTFVREAHKRGIEVILDVVFNHTA-EGN---ENGPILSFKGVDNTTYYMLAPKGEF 353
Db    169 ------FIDACHARRLGVLIDAVFNHLGPSGNYLPRFGPYLS--SASN-------PWGDG 213

Qy    354 YNYSGCGNTFNCNHPVVRQFIVDCLRYWVTEMHVDGFRFDLASIMTRGSSLWDPVNVYGA 413
Db    214 INIAG------ADSDEVRHYIIDCALRWMRDFHADGLRLDAVHALVDTTAVHVLEELANA 267

Qy    414 P--IEGDMITTGTPLVTPPLIDMISNDPILGGVKLIAEAWDAGGLYQVGQFPHWNVWSEW 471
Db    268 TRWLSGQL---GRPLSL--IAETDRNDP-----RLITRPSHGG----------YGITAQW 307

Qy    472 NGKYRDIVRQFIKGTDGFAGGFAECLCGSPHLYQAGGRKPWHSINFVCAHDGFTLADLVT 531
Db    308 N----DDIHHAI-----------------HTAVSGERQGYYA-------DFGSLATLA- 337

Qy    532 YNKKYNLPNG----------ENNRDGE---------NHNLSWNCGEEGEFARLSVKRLRK 572
Db    338 ----YTLRNGYFHAGTYSSFRRRRHGRALDTSAIPATRLLAYTCTHD----QVGNRALGD 389

Qy    573 RQMR-----NFFVCLMVSQGVP---MFYMGDEYGHTKGGNNNTYC--------HDSYVNY 616
Db    390 RPSQYLTGGQLAIKAALTLGSPYTAMLFMGEEWG--ASSPFQFFCSHPEPELAHSTVAG- 446

Qy    617 FRWDKKEQYSELHRFCCLMTKFRKECEGLGLEDFPTAK--------RLQWHGHQPGK--- 665
Db    447 ----RKEEFAE--------------HGWAADDIPDPQDPQTFQRCKLNWAEAGSGEHAR 487

Qy    666 --------------------------PDWSENSRFVAFSMKDERQGEIYVAFNTSHLPA 698
Db    488 LHRFYRDLIALRHNEADLADPWLDHLMVDYDEQQRWVVM-----RRGQLMIACNLGAEPT 542

Qy    699 VVELPERAGRRWE-PVV 714
Db    543 CVPVSGELVLAWESPII 559
```

```
RESULT  14
PULA_KLEPN
ID   PULA_KLEPN     STANDARD;      PRT;  1090 AA.
AC   P07206;
DT   01-APR-1988 (Rel. 07, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Pullulanase precursor (EC 3.2.1.41) (Alpha-dextrin endo-1,6-alpha-
DE   glucosidase) (Pullulan 6-glucanohydrolase).
GN   PULA.
OS   Klebsiella pneumoniae.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Klebsiella.
OX   NCBI_TaxID=573;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=UNF 5023;
RX   MEDLINE=90205629; PubMed=2181242;
RA   Kornacker M.G., Pugsley A.P.;
RT   "Molecular characterization of pulA and its product, pullulanase, a
RT   secreted enzyme of Klebsiella pneumoniae UNF5023.";
RL   Mol. Microbiol. 4:73-85(1990).
RN   [2]
RP   SEQUENCE OF 1-62 FROM N.A.
RX   MEDLINE=86033621; PubMed=3902792;
RA   Chapon C., Raibaud O.;
RT   "Structure of two divergent promoters located in front of the gene
RT   encoding pullulanase in Klebsiella pneumoniae and positively
RT   regulated by the malT product.";
RL   J. Bacteriol. 164:639-645(1985).
RN   [3]
RP   SEQUENCE OF 944-1090 FROM N.A.
RX   MEDLINE=89291709; PubMed=2661532;
RA   D'Enfert C., Pugsley A.P.;
RT   "Klebsiella pneumoniae pulS gene encodes an outer membrane
RT   lipoprotein required for pullulanase secretion.";
RL   J. Bacteriol. 171:3673-3679(1989).
CC   -!- CATALYTIC ACTIVITY: Hydrolysis of (1->6)-alpha-D-glucosidic
CC       linkages in pullulan and in amylopectin and glycogen, and the
CC       alpha-and beta-limit dextrins of amylopectin and glycogen.
CC   -!- SUBUNIT: HOMOTRIMER.
CC   -!- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
CC       (Probable).
CC   -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC       KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X52181; CAA36431.1; -.
DR   EMBL; M12503; AAA25087.2; -.
DR   EMBL; M29097; AAA61976.1; -.
DR   PIR; A25025; A25025.
DR   PIR; A32880; A32880.
DR   PIR; S11823; S11823.
DR   InterPro; IPR000461; Alpha_amylase.
DR   InterPro; IPR004193; isoamylase_N.
DR   Pfam; PF00128; alpha-amylase; 1.
DR   Pfam; PF02922; isoamylase_N; 1.
DR   PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW   Hydrolase; Glycosidase; Membrane; Lipoprotein; Signal.
FT   SIGNAL        1     19
FT   CHAIN        20   1090       PULLULANASE.
FT   LIPID        20     20       N-ACYL DIGLYCERIDE.
FT   ACT_SITE    684    684       BY SIMILARITY.
FT   ACT_SITE    713    713       BY SIMILARITY.
FT   ACT_SITE    841    841       BY SIMILARITY.
FT   CONFLICT      6      6       R -> C (IN REF. 2).
FT   CONFLICT     10     10       V -> F (IN REF. 2).
FT   CONFLICT     15     15       V -> I (IN REF. 2).
FT   CONFLICT     23     23       G -> S (IN REF. 2).
FT   CONFLICT     31     31       N -> S (IN REF. 2).
FT   CONFLICT     34     34       T -> N (IN REF. 2).
FT   CONFLICT     36     36       D -> DGNP (IN REF. 2).
FT   CONFLICT     55     58       TAVE -> MATA (IN REF. 2).
SQ   SEQUENCE   1090 AA;  118098 MW;  240AE7DFB3FF1BD6 CRC64;
```

```
  Query Match             6.3%;  Score 265;  DB 1;  Length 1090;
  Best Local Similarity   19.7%;  Pred. No. 3.4e-12;
  Matches  151;  Conservative   89;  Mismatches  244;  Indels  284;  Gaps   35;

Qy     21 GEVCA------AVVEAATKVEDEGEEDEPVAEDRYALGGACRVLAGMPAPLGATALAGGV 74
Db    266 GETVAIAAAEDGILISATQVQTAGVLDDAYAEAAEALS-----------YGAQLADGGV 313

Qy     75 NFAVYSGGATAAALCLFTPEDLKADRVTEEVPLDPLMNRTGNVWHVFIEGELHNMLYGYR 134
Db    314 TFRVWAPTAQQVDVVVYSAD----KKVIGSHPM--TRDSASGAWSWQGGSDLKGAFYRYA 367

Qy    135 FDGTFAPHCGHYLDVSNVVVDPYAKAVISRGEYGVPARGNNCWPQMAGMIPLPYSTFDWE 194
Db    368 MT-VYHPQSRKVEQYE--VTDPYAHSLSTNSEYS----------QVVDLNDSALKPDGWD 414

Qy    195 GDLPLRYPQK------DLVIYEMHLRGFTKHDSS-NVEHPGTFI----GAVSKLDYLKEL 243
```

```
Db    415 -NLTMPHAQKTKADLAKMTIHESHIRDLSAWDQTVPAELRGKYLALTAGDSNMVQHLKTL 473

Qy    244 ---GVNCIELMP------CHEF------------------------------------ 256
Db    474 SASGVTHVELLPVFDLATVNEFSDKVADIQQPFSRLCEVNSAVKSSEFAGYCDSGSTVEE 533

Qy    257 --NELEYSTSSSKMNF-------------WGYSTINFFSPMTRYTSGGIKNCGRDAINE 300
Db    534 VLNQLKQSDSQDNPQVQALNTLVAQTDSYNWGYDPFHYTVPEGSYATDPE---GTTRIKE 590

Qy    301 FKTFVREAHKR--GIEVILDVVFNHTAEGNENGPILSFKGVDNTT--------YYMLAPK 350
Db    591 FRTMI-QAIKQDLGMNVIMDVVYNHT---NAAGP------TDRTSVLDKIVPWYYQ---- 636

Qy    351 GEFYNYSGCGNTFNC------NHPVVRQFIVDCLRYWVTEMHVDGFRFDLASIMTRGS-- 402
Db    637 -RLNETTGSVESATCCSDSAPEHRMFAKLIADSLAVWTTDYKIDGFRFDLMGYHPKAQIL 695

Qy    403 SLWDPVNVYGAPIEGDMITTGTPLVTPPLIDMISNDPILGGVKLIAEAWDAGGLYQVGQF 462
Db    696 SAWERIKALNPDI--------------------------YFFGEGWDS------NQS 720

Qy    463 PHWNVWSEWN------GKYRDIVRQFIKG----------------------------T 486
Db    721 DRFEIASQINLKGTGIGTFSDRLRDSVRGGGPFDSGDALRQNQGIGSGAGVLPNELASLS 780

Qy    487 D------------GFAGGFAECLC----------------GSPHLYQAGGRKPWHSINFV 518
Db    781 DDQVRHLADLTRLGMAGNLADFVMIDKDGAAKKGSEIDYNGAPGGYAA---DPTEVVNYV 837

Qy    519 CAHDGFTLADLVTYNKKYNLPNGENNRDGENHNLSWNCGEEGEFARLSVKRLRKRQMRNF 578
Db    838 SKHDNQTLWDMISY----------------------KASQEADLA------TRVRMQAVS 869

Qy    579 FVCLMVSQGVPMFYMGDEYGHTKGGNNNTYCHDSYVNYFRWDKKEQYS 626
Db    870 LATVMLGQGIAFDQQGSELLRSK-----SFTRDSY-DSGDWFNRVDYS 911


RESULT  15
NEPU_THEVU
ID   NEPU_THEVU     STANDARD;      PRT;   585 AA.
AC   Q08751;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Neopullulanase (EC 3.2.1.135) (Alpha-amylase II).
GN   TVAII.
OS   Thermoactinomyces vulgaris.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Thermoactinomyces.
OX   NCBI_TaxID=2026;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=R-47;
RX   MEDLINE=93222535; PubMed=7763540;
RA   Tonozuka T., Ohtsuka M., Mogi S.-I., Sakai H., Ohta T., Sakano Y.;
RT   "A neopullulanase-type alpha-amylase gene from Thermoactinomyces
RT   vulgaris R-47.";
RL   Biosci. Biotechnol. Biochem. 57:395-401(1993).
RN   [2]
RP   X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).
RC   STRAIN=R-47;
RX   MEDLINE=99241045; PubMed=10222200;
RA   Kamitori S., Kondo S., Okuyama K., Yokota T., Shimura Y., Tonozuka T.,
RA   Sakano Y.;
RT   "Crystal structure of Thermoactinomyces vulgaris R-47 alpha-amylase II
RT   (TVAII) hydrolyzing cyclodextrins and pullulan at 2.6-A resolution.";
RL   J. Mol. Biol. 287:907-921(1999).
CC   -!- FUNCTION: HYDROLYZES PULLULAN EFFICIENTLY BUT ONLY A SMALL AMOUNT
CC       OF STARCH. ENDOHYDROLYSIS OF 1,4-ALPHA-GLUCOSIDIC LINKAGES IN
CC       PULLULAN TO FORM PANOSE. CLEAVES ALSO (1-6)-ALPHA-GLUCOSIDIC
CC       LINKAGES TO FORM MALTOTRIOSE.
CC   -!- CATALYTIC ACTIVITY: Hydrolysis of pullulan to panose (6-alpha-D-
CC       glucosylmaltose).
CC   -!- SUBUNIT: MONOMER.
CC   -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC       KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D13178; BAA02473.1; -.
DR   PDB; 1BVZ; 02-MAR-99.
DR   InterPro; IPR000461; Alpha_amylase.
DR   InterPro; IPR004185; alpha-amylase_N.
DR   Pfam; PF00128; alpha-amylase; 1.
DR   Pfam; PF02903; alpha-amylase_N; 1.
KW   Hydrolase; Glycosidase; Carbohydrate metabolism; 3D-structure.
FT   ACT_SITE    325    325       BY SIMILARITY.
FT   ACT_SITE    354    354       BY SIMILARITY.
FT   ACT_SITE    421    421       BY SIMILARITY.
SQ   SEQUENCE   585 AA;  67467 MW;  E311813A05A7791A CRC64;


  Query Match             6.1%;  Score 255;  DB 1;  Length 585;
  Best Local Similarity   21.4%;  Pred. No. 8.5e-12;
  Matches  132;  Conservative   79;  Mismatches  201;  Indels  204;  Gaps   32;

Qy    152 VVVDPYAKAVISRGEYGVPARGNNCWPQMAGMIPLPY-------STFDWEGDLPLRYPQK 204
Db     85 LLTGPQGEAVYF-GETGFSAERSK-----AGVFQYAYIHRSEVFTTPEW---------AK 129

Qy    205 DLVIYEMHLRGFTKHDSSNVEHPGT-----------------FIGAVSKLDYLKELGVN 246
Db    130 EAVIYQIFPERFANGDPSN-DPPGTEQWAKDARPRHDSFYGGDLKGVIDRLPYLEELGVT 188

Qy    247 CIELMPCHEFNELEYSTSSSKMNFWGYSTINFFSPMTRYTSGGIKNCGRDAINEFKTFVR 306
Db    189 ALYFTP------IFASPSHHK-----YDTADYLAIDPQFGD----------LPTFRRLVD 227

Qy    307 EAHKRGIEVILDVVFNHTAEGNENGPILSFKGVDNTTYYMLAPKGEFYNYSG-------- 358
Db    228 EAHRRGIKIILDAVFNHAGD-----QFFAFRDV--------LQKGEQSRYKDWFFIEDFP 274

Qy    359 CGNTFNCNH----------------PVVRQFIVDCLRYWVTEMHVDGFRFDLASIMTRG 401
Db    275 VSKTSRTNYETFAVQVPAMPKLRTENPEVKEYLFDVARFWM-EQGIDGWRLDVANEVDH- 332

Qy    402 SSLWDPVNVYGAPIEGDMITTGTPLVTPPLIDMISNDPILGGVKLIAEAW-DAGGLYQVG 460
Db    333 -AFWREFR-------------------RLVKSLNPDAL-----IVGEIWHDASGWLMGD 366

Qy    461 QFPHWNVWSEWNGKYRDIVRQFIKGTDGFAGGFAECLCGSPHLY-QAGGRKPWHSINFVC 519
Db    367 QFD-----SVMNYLFRESVIRFFATGEIHAERFDAELTRARMLYPEQAAQGLW---NLLD 418

Qy    520 AHDGFTLADLVTYNKKYNLPNGENNRDGENHNLSWNC-GEEGEFARLSVKRLRKRQMRNF 578
Db    419 SHD-------------------------TERFLTSCGGNEAKF-RLAV----------- 440

Qy    579 FVCLMVSQGVPMFYMGDEYGHTKGGNNNTYCHDSYVNYFRWDKKEQYSELHRFCCLMTKF 638
Db    441 -LFQMTYLGTPLIYYGDEIG--MAGATDPDCRRPMI----WEEKEQNRGLFEFYKELIRL 493

Qy    639 RKECEGLGLEDFPTAKRLQWHGHQPGKPDWSENSRFVAF--SMKDERQGEIYVAFNT--S 694
Db    494 RHR-----LASLTRGNVRSWHA--------DKQANLYAFVRTVQDQHVG---VVLNNRGE 537

Qy    695 HLPAVVELPERAGRRW 710
```

Db 538 KQTVLLQVPESGGKTW 553

Search completed: July 31, 2002, 12:15:20
Job time: 249 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: July 31, 2002, 12:08:11 ; Search time 69.26 Seconds
(without alignments)
1908.289 Million cell updates/sec

Title: US-09-674-817A-2
Perfect score: 4212
Sequence: 1 SGPAPRLRRWRPNATAGKGV.....SNLYPMLSYSSVILVLRPDV 764

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_19:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3568.5	84.7	789	10	O22637	O22637 zea mays (m
2	3527.5	83.7	818	10	Q41742	Q41742 zea mays (m
3	3491	82.9	733	10	O80403	O80403 oryza sativ
4	3001	71.2	569	10	Q9XFG6	Q9xfg6 hordeum vul
5	2895.5	68.7	783	10	O04196	O04196 arabidopsis
6	1809	42.9	327	10	Q9XFG7	Q9xfg7 triticum ae
7	1638.5	38.9	707	16	P73608	P73608 synechocyst
8	1538.5	36.5	702	10	Q9M0S5	Q9m0s5 arabidopsis
9	1491.5	35.4	332	10	Q9XFG8	Q9xfg8 solanum tub
10	1388	33.0	746	16	P72691	P72691 synechocyst
11	1305.5	31.0	664	16	Q9Z8F5	Q9z8f5 chlamydia p
12	1300	30.9	666	16	O84046	O84046 chlamydia t
13	1268	30.1	720	16	Q9RXP5	Q9rxp5 deinococcus
14	1264	30.0	666	16	Q9PKZ6	Q9pkz6 chlamydia m
15	1251	29.7	715	2	Q9KY03	Q9ky03 streptomyce
16	1249	29.7	715	2	Q9X947	Q9x947 streptomyce
17	1247	29.6	716	17	Q973H3	Q973h3 sulfolobus
18	1229.5	29.2	713	1	O05152	O05152 sulfolobus
19	1226	29.1	823	2	Q9LC80	Q9lc80 arthrobacte
20	1211	28.8	718	17	P95868	P95868 sulfolobus
21	1192.5	28.3	702	2	Q9ADI1	Q9adi1 streptomyce
22	1156.5	27.5	782	2	Q9L1E9	Q9l1e9 streptomyce
23	1149.5	27.3	716	16	Q9I1V5	Q9i1v5 pseudomonas
24	1123.5	26.7	698	16	Q92U64	Q92u64 rhizobium m
25	996	23.6	662	16	Q9CN93	Q9cn93 pasteurella
26	927.5	22.0	656	16	Q9KKS1	Q9kks1 vibrio chol
27	916.5	21.8	672	16	Q985P0	Q985p0 rhizobium l
28	890	21.1	882	10	Q9ZVT2	Q9zvt2 arabidopsis
29	864.5	20.5	662	16	Q92M11	Q92m11 rhizobium m
30	825	19.6	762	2	Q9RM63	Q9rm63 myroides od
31	812	19.3	168	10	Q9SPT7	Q9spt7 hordeum vul
32	811.5	19.3	656	2	Q9EUT3	Q9eut3 rhizobium t
33	579	13.7	332	2	Q54729	Q54729 synechococc
34	527.5	12.5	718	16	O34587	O34587 bacillus su
35	519.5	12.3	720	16	Q97FP9	Q97fp9 clostridium
36	512	12.2	825	2	Q59319	Q59319 caldocellum
37	505.5	12.0	717	16	Q9K7U5	Q9k7u5 bacillus ha
38	505	12.0	718	2	O69008	O69008 thermus sp.
39	449.5	10.7	1938	2	P70983	P70983 bacillus sp
40	448.5	10.6	668	2	P71095	P71095 bacteroides
41	442	10.5	849	2	Q9XDB5	Q9xdb5 fervidobact
42	418	9.9	759	16	Q97QT1	Q97qt1 streptococc
43	408	9.7	1072	16	Q9K6N1	Q9k6n1 bacillus ha
44	406.5	9.7	1165	16	Q99XX8	Q99xx8 streptococc
45	399.5	9.5	1287	2	Q9F930	Q9f930 streptococc

ALIGNMENTS

RESULT 1
O22637
ID O22637 PRELIMINARY; PRT; 789 AA.
AC O22637;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE SU1 ISOAMYLASE.
GN SUGARY1.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RA Beatty M.K., Myers A.M., James M.G.;
RT "Genomic nucleotide sequence of a full-length wild-type allele of the
RT maize sugary1 (Su1) gene (Accession No. AF030882) (PGR97-187).";
RL Plant Physiol. 115:1731-1731(1997).
DR EMBL; AF030882; AAB97167.1; -.
DR HSSP; P10342; 1BF2.
DR InterPro; IPR000461; Alpha_amylase.
DR InterPro; IPR004193; isoamylase_N.
DR Pfam; PF00128; alpha-amylase; 1.
DR Pfam; PF02922; isoamylase_N; 1.
SQ SEQUENCE 789 AA; 88352 MW; F3046F120834EBC4 CRC64;

Query Match 84.7%; Score 3568.5; DB 10; Length 789;
Best Local Similarity 83.6%; Pred. No. 9.8e-277;
Matches 654; Conservative 44; Mismatches 65; Indels 19; Gaps 6;

```
Qy       1 SGPAPRLR----RW------RPNATAGKGVGEV------CAAVVEAATKVEDEGEEDEPV 44
           | | | |     ||      |||  || | | :         | ||    ||: ::|| |
Db       9 SSPRPLLAVPAGRWRAGVRGRPN-VAGLGRGRLSLHAAAARPVAEAVQAEEDDDDDDEEV 67

Qy      45 AEDRYALGGACRVLAGMPAPLGATALAGGVNFAVYSGGATAAALCLFTPEDLKADRVTEE 104
           ||:|:|||||||||||||||||||| |||||||| ||:||:|||| | ||||||||||
```

```
Db      68 AEERFALGGACRVLAGMPAPLGATALRGGVNFAVYSSGASAASLCLFAPGDLKADRVTEE 127

Qy     105 VPLDPLMNRTGNVWHVFIEG-ELHNMLYGYRFDGTFAPHCGHYLDVSNVVVDPYAKAVIS 163
           ||||||:|||||||||||| | :|| |||||||| ||| | | ||||||||||||||:|
Db     128 VPLDPLLNRTGNVWHVFIHGDQLHGMLYGYRFDGVFAPERGQYYDVSNVVVDPYAKAVVS 187

Qy     164 RGEYGVPARGNNCWPQMAGMIPLPYSTFDWEGDLPLRYPQKDLVIYEMHLRGFTKHDSSN 223
           |||||||| | :|||||||||||||: |||:||||| | ||||||||||||||||:||
Db     188 RGEYGVPAPGGSCWPQMAGMIPLPYNKFDWQGDLPLGYHQKDLVIYEMHLRGFTKHNSSK 247

Qy     224 VEHPGTFIGAVSKLDYLKELGVNCIELMPCHEFNELEYSTSSSKMNFWGYSTINFFSPMT 283
            :||||:||||||||:|||||||||||||||||||||| :|||||||||||||||||||
Db     248 TKHPGTYIGAVSKLDHLKELGVNCIELMPCHEFNELEYFSSSSKMNFWGYSTINFFSPMA 307

Qy     284 RYTSGGIKNCGRDAINEFKTFVREAHKRGIEVILDVVFNHTAEGNENGPILSFKGVDNTT 343
           ||:| ||:: |  |||||| ||||||||||||:|||||||||||| |||||:|:||:|
Db     308 RYSSSGIRDSGCGAINEFKAFVREAHKRGIEVIMDVVFNHTAEGNEKGPILSFRGIDNST 367

Qy     344 YYMLAPKGEFYNYSGCGNTFNCNHPVVRQFIVDCLRYWVTEMHVDGFRFDLASIMTRGSS 403
           |||||||||||||||||||||||||||:|||||||||||||||||||||||:||| |
Db     368 YYMLAPKGEFYNYSGCGNTFNCNHPVVREFIVDCLRYWVTEMHVDGFRFDLASILTRGCS 427

Qy     404 LWDPVNVYGAPIEGDMITTGTPLVTPPLIDMISNDPILGGVKLIAEAWDAGGLYQVGQFP 463
           |||||||||:|:||||||||||||| ||||||||||||| ||||||||||||||||||||
Db     428 LWDPVNVYGSPMEGDMITTGTPLVAPPLIDMISNDPILGNVKLIAEAWDAGGLYQVGQFP 487

Qy     464 HWNVWSEWNGKYRDIVRQFIKGTDGFAGGFAECLCGSPHLYQAGGRKPWHSINFVCAHDG 523
           |||||||||||||| |||||||||||| |||||||| ||||||||||||||||||||||
Db     488 HWNVWSEWNGKYRDTVRQFIKGTDGFAGAFAECLCGSPQLYQAGGRKPWHSINFVCAHDG 547

Qy     524 FTLADLVTYNKKYNLPNGENNRDGENHNLSWNCGEEGEFARLSVKRLRKRQMRNFFVCLM 583
           |||||||||| |||| |||:||||||||||||||||||||| |||:||||||||||||||
Db     548 FTLADLVTYNSKYNLSNGEDNRDGENHNLSWNCGEEGEFASLSVRRLRKRQMRNFFVCLM 607

Qy     584 VSQGVPMFYMGDEYGHTKGGNNNTYCHDSYVNYFRWDKK-EQYSELHRFCCLMTKFRKEC 642
           |||||||||||||||||||||||||||| |||||||||| || |:|:|||| |||||||||
Db     608 VSQGVPMFYMGDEYGHTKGGNNNTYCHDHYVNYFRWDKKEEQSSDLYRFCRLMTKFRKEC 667

Qy     643 EGLGLEDFPTAKRLQWHGHQPGKPDWSENSRFVAFSMKDERQGEIYVAFNTSHLPAVVEL 702
           | |||||||::||:|||||||||||| |||||:|||| :|||||||||||| || |
Db     668 ESLGLEDFPTSERLKWHGHQPGKPDWSEASRFVAFTMKDETKGEIYVAFNTSHLPVVVGL 727

Qy     703 PERAGRRWEPVVDTGKPAPYDFLTDDLPDRALTIHQFSHFLYSNLYPMLSYSSVILVLRP 762
           |||:| |||||||||| |||||||| |||||:|::|||||| ||||||||||||:|||||
Db     728 PERSGFRWEPVVDTGKEAPYDFLTDGLPDRAVTVYQFSHFLNSNLYPMLSYSSIILVLRP 787

Qy     763 DV 764
           ||
Db     788 DV 789


RESULT   2
Q41742
ID   Q41742      PRELIMINARY;      PRT;   818 AA.
AC   Q41742;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   SU1P (FRAGMENT).
GN   SUGARY1.
OS   Zea mays (Maize).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC   Panicoideae; Andropogoneae; Zea.
OX   NCBI_TaxID=4577;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=ENDOSPERM;
RX   MEDLINE=95290995; PubMed=7773016;
RA   James M.G., Robertson D.S., Myers A.M.;
RT   "Characterization of the maize gene sugary1, a determinant of starch
RT   composition in kernels.";
RL   Plant Cell 7:417-429(1995).
DR   EMBL; U18908; AAA91298.1; -.
DR   HSSP; P10342; 1BF2.
DR   InterPro; IPR000461; Alpha_amylase.
DR   InterPro; IPR004193; isoamylase_N.
DR   Pfam; PF00128; alpha-amylase; 1.
DR   Pfam; PF02922; isoamylase_N; 1.
FT   NON_TER       1       1
SQ   SEQUENCE   818 AA;  91535 MW;  4779CA8508B1AFED CRC64;


  Query Match             83.7%;  Score 3527.5;  DB 10;  Length 818;
  Best Local Similarity   83.0%;  Pred. No. 2e-273;
  Matches  649;  Conservative   44;  Mismatches   70;  Indels   19;  Gaps    6;

Qy       1 SGPAPRLR----RW------RPNATAGKGVGEV------CAAVVEAATKVEDEGEEDEPV 44
           | | | |      ||       |||  || | | :          | ||    ||: ::|| |
Db      38 SSPRPLLAVPAGRWRAGVRGRPN-VAGLGRGRLSLHAAAARPVAEAVQAEEDDDDDDEEV 96

Qy      45 AEDRYALGGACRVLAGMPAPLGATALAGGVNFAVYSGGATAAALCLFTPEDLKADRVTEE 104
           ||:|:|||||||||||||||||| |||||||| ||:||:| || | ||||||||||
Db      97 AEERFALGGACRVLAGMPAPLGATALRGGVNFAVYSSGASAASLSLFAPGDLKADRVTEE 156

Qy     105 VPLDPLMNRTGNVWHVFIEG-ELHNMLYGYRFDGTFAPHCGHYLDVSNVVVDPYAKAVIS 163
           ||||||:|||||||||||| | ||| || |||||| |||  | | ||||||||||||:|
Db     157 VPLDPLLNRTGNVWHVFIHGDELHGMLCGYRFDGVFAPERGQYYDVSNVVVDPYAKAVVS 216

Qy     164 RGEYGVPARGNNCWPQMAGMIPLPYSTFDWEGDLPLRYPQKDLVIYEMHLRGFTKHDSSN 223
           |||||||| | :|||||||||||||: |||:||||| | ||||||||||||||||:||
Db     217 RGEYGVPAPGGSCWPQMAGMIPLPYNKFDWQGDLPLGYHQKDLVIYEMHLRGFTKHNSSK 276

Qy     224 VEHPGTFIGAVSKLDYLKELGVNCIELMPCHEFNELEYSTSSSKMNFWGYSTINFFSPMT 283
            :||||:||||||||:|||||||||||||||||||||| :|||||||||||||||||||
Db     277 TKHPGTYIGAVSKLDHLKELGVNCIELMPCHEFNELEYFSSSSKMNFWGYSTINFFSPMA 336

Qy     284 RYTSGGIKNCGRDAINEFKTFVREAHKRGIEVILDVVFNHTAEGNENGPILSFKGVDNTT 343
           ||:| ||:: |  |||||| ||||||||||||:|||||||||||| |||||:|:||:|
Db     337 RYSSSGIRDSGCGAINEFKAFVREAHKRGIEVIMDVVFNHTAEGNEKGPILSFRGIDNST 396

Qy     344 YYMLAPKGEFYNYSGCGNTFNCNHPVVRQFIVDCLRYWVTEMHVDGFRFDLASIMTRGSS 403
           |||||||||||||||||||||||||||:|||||||||||||||||||||||:||| |
Db     397 YYMLAPKGEFYNYSGCGNTFNCNHPVVREFIVDCLRYWVTEMHVDGFRFDLASILTRGCS 456

Qy     404 LWDPVNVYGAPIEGDMITTGTPLVTPPLIDMISNDPILGGVKLIAEAWDAGGLYQVGQFP 463
           |||||||||:|:||||||||||||| ||||||||||||| ||||||||||||||| ||||
Db     457 LWDPVNVYGSPMEGDMITTGTPLVAPPLIDMISNDPILGNVKLIAEAWDAGGLYQEGQFP 516

Qy     464 HWNVWSEWNGKYRDIVRQFIKGTDGFAGGFAECLCGSPHLYQAGGRKPWHSINFVCAHDG 523
           |||||||||||||| |||||||||||| |||||||| ||||||||||||||||||||||
Db     517 HWNVWSEWNGKYRDTVRQFIKGTDGFAGAFAECLCGSPQLYQAGGRKPWHSIGFVCAHDG 576

Qy     524 FTLADLVTYNKKYNLPNGENNRDGENHNLSWNCGEEGEFARLSVKRLRKRQMRNFFVCLM 583
           |||||||||| |||| |||: |||||||||||||||||||| |||:|||||||||||||
Db     577 FTLADLVTYNSKYNLSNGEDFRDGENHNLSWNCGEEGEFASLSVRRLRKRQMRNFFVCLM 636

Qy     584 VSQGVPMFYMGDEYGHTKGGNNNTYCHDSYVNYFRWDKK-EQYSELHRFCCLMTKFRKEC 642
           |||||||||||||||||||||||||||| |||||||||| || |:|:|||| |||:|||||
Db     637 VSQGVPMFYMGDEYGHTKGGNNNTYCHDHYVNYFRWDKKEEQSSDLYRFCRLMTEFRKEC 696

Qy     643 EGLGLEDFPTAKRLQWHGHQPGKPDWSENSRFVAFSMKDERQGEIYVAFNTSHLPAVVEL 702
           | |||||||::||:|||||||||||| |||||:|||| :|||||||||||| || |
Db     697 ESLGLEDFPTSERLKWHGHQPGKPDWSEASRFVAFTMKDETKGEIYVAFNTSHLPVVVGL 756

Qy     703 PERAGRRWEPVVDTGKPAPYDFLTDDLPDRALTIHQFSHFLYSNLYPMLSYSSVILVLRP 762
           |||:| |||||||||| |||||||| |||||:|::|||||| ||||||||||||:|||||
Db     757 PERSGFRWEPVVDTGKEAPYDFLTDGLPDRAVTVYQFSHFLNSNLYPMLSYSSIILVLRP 816

Qy     763 DV 764
           ||
Db     817 DV 818


RESULT   3
O80403
```

```
ID   O80403      PRELIMINARY;      PRT;   733 AA.
AC   O80403;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   ISOAMYLASE (FRAGMENT).
OS   Oryza sativa (Rice).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=4530;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. FUJIHIKARI; TISSUE=ENDOSPERM;
RX   MEDLINE=99266133; PubMed=10333591;
RA   Fujita N., Kubo A., Francisco P.B., Nakakita M., Harada K., Minaka N.,
RA   Nakamura Y.;
RT   "Purification, characterization, and cDNA structure of isoamylase from
RT   developing endosperm of rice.";
RL   Planta 208:283-293(1999).
DR   EMBL; AB015615; BAA29041.1; -.
DR   HSSP; P10342; 1BF2.
DR   InterPro; IPR000461; Alpha_amylase.
DR   InterPro; IPR004193; isoamylase_N.
DR   Pfam; PF00128; alpha-amylase; 1.
DR   Pfam; PF02922; isoamylase_N; 1.
FT   NON_TER       1       1
SQ   SEQUENCE   733 AA;  82119 MW;  FC864DDDF0E30D35 CRC64;
```

```
  Query Match             82.9%;  Score 3491;  DB 10;  Length 733;
  Best Local Similarity   90.1%;  Pred. No. 1.4e-270;
  Matches  621;  Conservative   28;  Mismatches   40;  Indels    0;  Gaps    0;

Qy      44 VAEDRYALGGACRVLAGMPAPLGATALAGGVNFAVYSGGATAAALCLFTPEDLKADRVTE 103
           |  :||||||||||||||||||||||| |||||||| ||:||:|||||:||:|| |||
Db       5 VMPERYALGGACRVLAGMPAPLGATALDGGVNFAVYSAGASAASLCLFTPDDLEADEVTE 64

Qy     104 EVPLDPLMNRTGNVWHVFIEGELHNMLYGYRFDGTFAPHCGHYLDVSNVVVDPYAKAVIS 163
           ||||||| |||||||||||||||||||||||| |||||| | ||||||||||||||||
Db      65 EVPLDPLFNRTGNVWHVFIEGELHNMLYGYRFDGMFAPHCGQYFDVSNVVVDPYAKAVIS 124

Qy     164 RGEYGVPARGNNCWPQMAGMIPLPYSTFDWEGDLPLRYPQKDLVIYEMHLRGFTKHDSSN 223
           |||||||  | :|||||||||||||||||:|||||||||||||||||||||||| |||
Db     125 RGEYGVPGPGGDCWPQMAGMIPLPYSTFDWQGDLPLRYPQKDLVIYEMHLRGFTKHSSSN 184

Qy     224 VEHPGTFIGAVSKLDYLKELGVNCIELMPCHEFNELEYSTSSSKMNFWGYSTINFFSPMT 283
           ||||||:|||:|||||||||||||:|||||||||||||| : ||||||||||||||||||
Db     185 VEHPGTYIGAISKLDYLKELGVNCVELMPCHEFNELEYFSCSSKMNFWGYSTINFFSPMI 244

Qy     284 RYTSGGIKNCGRDAINEFKTFVREAHKRGIEVILDVVFNHTAEGNENGPILSFKGVDNTT 343
           ||:||||:|||||||||||||||||||||:|||||||||||| |||||:|:||:|
Db     245 RYSSGGIRNCGRDAINEFKTFVREAHKRGIEVIMDVVFNHTAEGNEKGPILSFRGIDNST 304

Qy     344 YYMLAPKGEFYNYSGCGNTFNCNHPVVRQFIVDCLRYWVTEMHVDGFRFDLASIMTRGSS 403
           |||||||||||||||||||||||||:|||||||||||||||||||||||||||| |
Db     305 YYMLAPKGEFYNYSGCGNTFNCNHPVVREFIVDCLRYWVTEMHVDGFRFDLASIMTRGCS 364

Qy     404 LWDPVNVYGAPIEGDMITTGTPLVTPPLIDMISNDPILGGVKLIAEAWDAGGLYQVGQFP 463
           |||||||||:|:|||| |||||| |||||||||||||| ||||||||||||||||||||
Db     365 LWDPVNVYGSPVEGDMTTTGTPLATPPLIDMISNDPILGDVKLIAEAWDAGGLYQVGQFP 424

Qy     464 HWNVWSEWNGKYRDIVRQFIKGTDGFAGGFAECLCGSPHLYQAGGRKPWHSINFVCAHDG 523
           || :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     425 HWKIWSEWNGKYRDIVRQFIKGTDGFAGGFAECLCGSPHLYQAGGRKPWHSINFVCAHDG 484

Qy     524 FTLADLVTYNKKYNLPNGENNRDGENHNLSWNCGEEGEFARLSVKRLRKRQMRNFFVCLM 583
           |||||||||||||  |||:|||||||||||||||||| |||||||||||||||  ||
Db     485 FTLADLVTYNKKYNSSNGEDNRDGENHNLSWNCGEEGEFAGLSVKRLRKRQMRNFFVSLM 544

Qy     584 VSQGVPMFYMGDEYGHTKGGNNNTYCHDSYVNYFRWDKKEQYSELHRFCCLMTKFRKECE 643
           |||||||||||||||||||||||||| |||||||||: |:| ||| ;|||||:||
Db     545 VSQGVPMFYMGDEYGHTKGGNNNTYCHDHYVNYFRWDKKEESSDLQRFCSLMTKFRKQCE 604

Qy     644 GLGLEDFPTAKRLQWHGHQPGKPDWSENSRFVAFSMKDERQGEIYVAFNTSHLPAVVELP 703
           ||| ||||||:|| |||||||||||||| ||||||| ||| :|||||||| ||||||| ||
Db     605 SLGLADFPTAQRLHWHGHQPGKPDWSETSRFVAFSTKDETKGEIYVAFNASHLPAVVGLP 664

Qy     704 ERAGRRWEPVVDTGKPAPYDFLTDDLPDR 732
           || | ||||:|||||||||||||||||||
Db     665 ERPGYRWEPLVDTGKPAPYDFLTDDLPDR 693
```

```
RESULT   4
Q9XFG6
ID   Q9XFG6      PRELIMINARY;      PRT;   569 AA.
AC   Q9XFG6;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   ISOAMYLASE 1 (FRAGMENT).
OS   Hordeum vulgare (Barley).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC   Triticeae; Hordeum.
OX   NCBI_TaxID=4513;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. BOMI;
RX   MEDLINE=99364538; PubMed=10437827;
RA   Sun C., Sathish P., Ahlandsberg S., Jansson C.;
RT   "Analyses of isoamylase gene activity in wild-type barley indicate its
RT   involvement in starch synthesis.";
RL   Plant Mol. Biol. 40:431-443(1999).
DR   EMBL; AF142589; AAD33889.1; -.
DR   HSSP; P10342; 1BF2.
DR   InterPro; IPR000461; Alpha_amylase.
DR   InterPro; IPR004193; isoamylase_N.
DR   Pfam; PF00128; alpha-amylase; 1.
DR   Pfam; PF02922; isoamylase_N; 1.
FT   NON_TER       1       1
FT   NON_TER     569     569
SQ   SEQUENCE   569 AA;  64020 MW;  5E5B8B12427D435A CRC64;
```

```
  Query Match             71.2%;  Score 3001;  DB 10;  Length 569;
  Best Local Similarity   94.7%;  Pred. No. 1.5e-231;
  Matches  541;  Conservative   12;  Mismatches   16;  Indels    2;  Gaps    2;

Qy      51 LGGACRVLAGMPAPLGATALAGGVNFAVYSGGATAAALCLFTPEDLKADRVTEEVPLDPL 110
           |||||||||||||  |||||||||||||||      |||||||||||||:||||||||
Db       1 LGGACRVLAGMPA-AGATALAGGVNFAVYS-AEPPRALCLFTPEDLKADRVSEEVPLDPL 58

Qy     111 MNRTGNVWHVFIEGELHNMLYGYRFDGTFAPHCGHYLDVSNVVVDPYAKAVISRGEYGVP 170
           |||||:|||||:|||| |||||||||||||||| |||||||||||||||| |||||
Db      59 MNRTGDVWHVFLEGELHGMLYGYRFDGTFAPHCGHYFDVSNVVVDPYAKAVISREEYGVP 118

Qy     171 ARGNNCWPQMAGMIPLPYSTFDWEGDLPLRYPQKDLVIYEMHLRGFTKHDSSNVEHPGTF 230
           | ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     119 AHGNNSWPQMAGMIPLPYSTFDWEGDLPLRYPQKDLVIYEMHLRGFTKHDSSNVEHPGTF 178

Qy     231 IGAVSKLDYLKELGVNCIELMPCHEFNELEYSTSSSKMNFWGYSTINFFSPMTRYTSGGI 290
           |||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db     179 IGAVSKLDYLKELGVNCIELMPCHEFNELEYATSSSKMNFWGYSTINFFSPMTRYTSGGI 238

Qy     291 KNCGRDAINEFKTFVREAHKRGIEVILDVVFNHTAEGNENGPILSFKGVDNTTYYMLAPK 350
           |||||| ||||||||||:|||||||||||||||||||||||||||:||||||||||||
Db     239 KNCGRDGINEFKTFVRESHKRGIEVILDVVFNHTAEGNENGPILSFRGVDNTTYYMLAPK 298

Qy     351 GEFYNYSGCGNTFNCNHPVVRQFIVDCLRYWVTEMHVDGFRFDLASIMTRGSSLWDPVNV 410
           |||||||||||||||||||||||||||||||| |||:|||||||||||||||||||:||
Db     299 GEFYNYSGCGNTFNCNHPVVRQFIVDCLRYWVMEMHIDGFRFDLASIMTRGSSLWDPINV 358

Qy     411 YGAPIEGDMITTGTPLVTPPLIDMISNDPILGGVKLIAEAWDAGGLYQVGQFPHWNVWSE 470
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     359 YGAPIEGDMITTGTPLVTPPLIDMISNDPILGGVKLIAEAWDAGGLYQVGQFPHWNVWSE 418
```

```
Qy     471 WNGKYRDIVRQFIKGTDGFAGGFAECLCGSPHLYQAGGRKPWHSINFVCAHDGFTLADLV 530
           |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db     419 WNGKYRDIVRQFIKGTDGFAGGFAECLCGSPQLYQAGGRKPWHSINFVCAHDGFTLADLV 478

Qy     531 TYNKKYNLPNGENNRDGENHNLSWNCGEEGEFARLSVKRLRKRQMRNFFVCLMVSQGVPM 590
           ||: |||||||:|||||||||||||||||||| ||||||||||||||||||||||||||:
Db     479 TYHNKYNLPNGEDNRDGENHNLSWNCGEEGEFARSSVKRLRKRQMRNFFVCLMVSQGVPI 538

Qy     591 FYMGDEYGHTKGGNNNTYCHDSYVNYFRWDK 621
           |||||||||||||||||||||||||||||:|
Db     539 FYMGDEYGHTKGGNNNTYCHDSYVNYFRWEK 569


RESULT   5
O04196
ID   O04196      PRELIMINARY;      PRT;   783 AA.
AC   O04196;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PUTATIVE ISOAMYLASE.
GN   AT2G39930.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RX   MEDLINE=20083487; PubMed=10617197;
RA   Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA   Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
RA   Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA   Cronin L.A., Shen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.E.,
RA   Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA   Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
RA   Salzberg S.L., Fraser C.M., Venter J.C.;
RT   "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT   thaliana.";
RL   Nature 402:761-768(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RA   Lin X.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF002109; AAB95278.1; -.
DR   HSSP; P10342; 1BF2.
DR   InterPro; IPR000461; Alpha_amylase.
DR   InterPro; IPR004193; isoamylase_N.
DR   Pfam; PF00128; alpha-amylase; 1.
DR   Pfam; PF02922; isoamylase_N; 1.
SQ   SEQUENCE   783 AA;  89480 MW;  638A6F9088CDB896 CRC64;


  Query Match             68.7%;  Score 2895.5;  DB 10;  Length 783;
  Best Local Similarity   69.3%;  Pred. No. 7e-223;
  Matches  516;  Conservative   84;  Mismatches  128;  Indels   17;  Gaps    5;

Qy      28 VEAATKVEDEGEE----DEPVAEDRYALGGACRVLAGMPAPLGATALAGGVNFAVYSGGA 83
           : |  :  :| |     ::|:  ||: :        |:|:| | |    ||||:|||  :
Db      42 ISAKDRRSNEAENIAVVEKPLKSDRFFISD------GLPSPFGPTVRDDGVNFSVYSTNS 95

Qy      84 TAAALCLFTPEDLKADRVTEEVPLDPLMNRTGNVWHVFIEGELHNMLYGYRFDGTFAPHC 143
             :| :|| :  ||: ::|||| ||| ||||:|||||: |:  :|||||||| |:|
Db      96 VSATICLISLSDLRQNKVTEEIQLDPSRNRTGHVWHVFLRGDFKDMLYGYRFDGKFSPEE 155

Qy     144 GHYLDVSNVVVDPYAKAVISRGEYGVPARGNNCWPQMAGMIPLPYSTFDWEGDLPLRYPQ 203
           ||| | ||:::|||||:||| |:||    :|||||| |:|     ||||||: |: ||
Db     156 GHYYDSSNILLDPYAKAIISRDEFGVLGPDDNCWPQMACMVPTREEEFDWEGDMHLKLPQ 215

Qy     204 KDLVIYEMHLRGFTKHDSSNVEHPGTFIGAVSKLDYLKELGVNCIELMPCHEFNELEYST 263
           |||||||||:||||:|:|| :| |||: |   |||:|||||:|||||||||||||||| :
Db     216 KDLVIYEMHVRGFTRHESSKIEFPGTYQGVAEKLDHLKELGINCIELMPCHEFNELEYYS 275

Qy     264 SSS-----KMNFWGYSTINFFSPMTRYTSGGIKNCGRDAINEFKTFVREAHKRGIEVILD 318
           ::      ::|||||||| ||||| || |    |    ||||||  |:||||||||||:|
Db     276 YNTILGDHRVNFWGYSTIGFFSPMIRYASASSNNFAGRAINEFKILVKEAHKRGIEVIMD 335

Qy     319 VVFNHTAEGNENGPILSFKGVDNTTYYMLAPKGEFYNYSGCGNTFNCNHPVVRQFIVDCL 378
           || |||||||| ||| ||:|||: |||||||||||||||||||||||||||||:|||
Db     336 VVLNHTAEGNEKGPIFSFRGVDNSVYYMLAPKGEFYNYSGCGNTFNCNHPVVRQFILDCL 395

Qy     379 RYWVTEMHVDGFRFDLASIMTRGSSLWDPVNVYGAPIEGDMITTGTPLVTPPLIDMISND 438
           |||||||||||||||| |||:| |||||  ||||| :|||:|||||:  ||:|||||||
Db     396 RYWVTEMHVDGFRFDLGSIMSRSSSLWDAANVYGADVEGDLLTTGTPISCPPVIDMISND 455

Qy     439 PILGGVKLIAEAWDAGGLYQVGQFPHWNVWSEWNGKYRDIVRQFIKGTDGFAGGFAECLC 498
           ||| |||||||||||||||| |||| :|||||||:||:|||||||||||:| ||||||
Db     456 PILRGVKLIAEAWDAGGLYQVGMFPHWGIWSEWNGKFRDVVRQFIKGTDGFSGAFAECLC 515

Qy     499 GSPHLYQAGGRKPWHSINFVCAHDGFTLADLVTYNKKYNLPNGENNRDGENHNLSWNCGE 558
           |||:||| ||||||||||:|||||||||||||| | || ||| | |||||| ||||||
Db     516 GSPNLYQ-GGRKPWHSINFICAHDGFTLADLVTYNNKNNLANGEENNDGENHNYSWNCGE 574

Qy     559 EGEFARLSVKRLRKRQMRNFFVCLMVSQGVPMFYMGDEYGHTKGGNNNTYCHDSYVNYFR 618
           ||:|| :|||||||||||||| ||||||||| ||||||||||||||||||||:|:||||
Db     575 EGDFASISVKRLRKRQMRNFFVSLMVSQGVPMIYMGDEYGHTKGGNNNTYCHDNYMNYFR 634

Qy     619 WDKKEQ-YSELHRFCCLMTKFRKECEGLGLEDFPTAKRLQWHGHQPGKPDWSENSRFVAF 677
           |||||: :|:  ||| :: ||| ||| ||| |||||||||||  |  |:||| ||||||
Db     635 WDKKEEAHSDFFRFCRILIKFRDECESLGLNDFPTAKRLQWHGLAPEIPNWSETSRFVAF 694

Qy     678 SMKDERQGEIYVAFNTSHLPAVVELPERAGRRWEPVVDTGKPAPYDFLTDDLPDRALTIH 737
           |: |  : |||||||||||  :| || | | |||| ||| ||:||| :| |||:|   :
Db     695 SLVDSVKKEIYVAFNTSHLATLVSLPNRPGYRWEPFVDTSKPSPYDCITPDLPERETAMK 754

Qy     738 QFSHFLYSNLYPMLSYSSVILVLRP 762
           |: ||| :|:|||||||||:||:| |
Db     755 QYRHFLDANVYPMLSYSSIILLLSP 779


RESULT   6
Q9XFG7
ID   Q9XFG7      PRELIMINARY;      PRT;   327 AA.
AC   Q9XFG7;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   ISOAMYLASE 1 (FRAGMENT).
OS   Triticum aestivum (Wheat).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC   Triticeae; Triticum.
OX   NCBI_TaxID=4565;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99364538; PubMed=10437827;
RA   Sun C., Sathish P., Ahlandsberg S., Jansson C.;
RT   "Analyses of isoamylase gene activity in wild-type barley indicate its
RT   involvement in starch synthesis.";
RL   Plant Mol. Biol. 40:431-443(1999).
DR   EMBL; AF142590; AAD33890.1; -.
DR   HSSP; P10342; 1BF2.
DR   InterPro; IPR000461; Alpha_amylase.
DR   Pfam; PF00128; alpha-amylase; 1.
DR   PRINTS; PR00110; ALPHAAMYLASE.
FT   NON_TER       1      1
FT   NON_TER     327    327
SQ   SEQUENCE   327 AA;  36876 MW;  9474EB0BEF4F07B9 CRC64;


  Query Match             42.9%;  Score 1809;  DB 10;  Length 327;
  Best Local Similarity   99.4%;  Pred. No. 1.6e-136;
  Matches  325;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy    196 DLPLRYPQKDLVIYEMHLRGFTKHDSSNVEHPGTFIGAVSKLDYLKELGVNCIELMPCHE 255
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 DLPLRYPQKDLVIYEMHLRGFTKHDSSNVEHPGTFIGAVSKLDYLKELGVNCIELMPCHE 60

Qy    256 FNELEYSTSSSKMNFWGYSTINFFSPMTRYTSGGIKNCGRDAINEFKTFVREAHKRGIEV 315
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 FNELEYSTSSSKMNFWGYSTINFFSPMTRYTSGGIKNCGRDAINEFKTFVREAHKRGIEV 120

Qy    316 ILDVVFNHTAEGNENGPILSFKGVDNTTYYMLAPKGEFYNYSGCGNTFNCNHPVVRQFIV 375
          ||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db    121 ILDVVFNHTAEGNENGPILSFRGVDNTTYYMLAPKGEFYNYSGCGNTFNCNHPVVRQFIV 180

Qy    376 DCLRYWVTEMHVDGFRFDLASIMTRGSSLWDPVNVYGAPIEGDMITTGTPLVTPPLIDMI 435
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 DCLRYWVTEMHVDGFRFDLASIMTRGSSLWDPVNVYGAPIEGDMITTGTPLVTPPLIDMI 240

Qy    436 SNDPILGGVKLIAEAWDAGGLYQVGQFPHWNVWSEWNGKYRDIVRQFIKGTDGFAGGFAE 495
          ||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 SNDPILGGVKLVAEAWDAGGLYQVGQFPHWNVWSEWNGKYRDIVRQFIKGTDGFAGGFAE 300

Qy    496 CLCGSPHLYQAGGRKPWHSINFVCAHD 522
          |||||||||||||||||||||||||||
Db    301 CLCGSPHLYQAGGRKPWHSINFVCAHD 327
```

```
RESULT   7
P73608
ID   P73608      PRELIMINARY;      PRT;   707 AA.
AC   P73608;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   GLYCOGEN OPERON PROTEIN GLGX.
GN   GLGX OR SLR1857.
OS   Synechocystis sp. (strain PCC 6803).
OC   Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX   NCBI_TaxID=1148;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97061201; PubMed=8905231;
RA   Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA   Miyajima N., Hirosawa M., Sugiura M., Sasamoto S., Kimura T.,
RA   Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA   Shimpo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA   Tabata S.;
RT   "Sequence analysis of the genome of the unicellular cyanobacterium
RT   Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT   entire genome and assignment of potential protein-coding regions.";
RL   DNA Res. 3:109-136(1996).
DR   EMBL; D90908; BAA17652.1; -.
DR   HSSP; P10342; 1BF2.
DR   InterPro; IPR000461; Alpha_amylase.
DR   InterPro; IPR004193; isoamylase_N.
DR   Pfam; PF00128; alpha-amylase; 1.
DR   Pfam; PF02922; isoamylase_N; 1.
KW   Complete proteome.
SQ   SEQUENCE   707 AA;  79895 MW;  CFA27B3C86318DB1 CRC64;
```

```
  Query Match             38.9%;  Score 1638.5;  DB 16;  Length 707;
  Best Local Similarity   49.1%;  Pred. No. 2.4e-122;
  Matches  342;  Conservative   77;  Mismatches  216;  Indels   61;  Gaps   16;

Qy     56 RVLAGMPAPLGATALAGGVNFAVYSGGATAAALCLFTPEDLKADRVTEEVPLDPLMNRTG 115
          ::  | | | ||| : |||||::||  :||  | ||   : :| :   |:|  |   | |
Db     16 KLRCGQPFPFPGATIVPGGVNFSIYSSHSTACTLVLF---EKRAPQPFVE1PF-PESFRIG 71

Qy    116 NVW-HVFIEGELHNMLYGYRFDGTFAPHCGHYLDVSNVVVDPYAKAVISRGEYGVPARGN 174
          ||:  |  : :  |: |||| :|      ||: | | |::||||| |  |  :|     :
Db     72 NVYCMVVFDLDFENLEYGYRMEGPNNFQQGHWFDPSKVLLDPYAKVVSGRDVWGTQPNWD 131

Qy    175 NCWPQMAGMIPLPYSTFDWEGDLPLRYPQKDLVIYEMHLRGFTKHDSSNVE--HPGTFIG 232
          : : |  |   | :  |||| | ||  | :|:||||||:|||||  || |:  | ||| |
Db    132 DIY-QHRGR--LSFDDFDWENDSPLDVPLEDMVIYEMHVRGFTKDPSSGVKENHRGTFAG 188

Qy    233 AVSKLDYLKELGVNCIELMPCHEFNELEYS-----TSSSKMNFWGYSTINFFSPMTRYTS 287
          :||: ||:|||| ||||  ||:| |:|     |    :|:||||:|||:|   | :
Db    189 ILSKIPYLQELGVNTIELMPIFEFDEFEHSRYHPETGEFLVNYWGYSTVNFFAPKAGYAA 248

Qy    288 GGIKNCGRDAINEFKTFVREAHKRGIEVILDVVFNHTAEGNENGPILSFKGVDNTTYYML 347
          |    |   |:| |  |:| || || |||||||||||||||| || :||:|:|| |||||
Db    249 TG--KFGMQ-IDELKNLVKELHKVGISVILDVVFNHTAEGNERGPTISFRGLDNKTYYML 305

Qy    348 APKGEFYNYSGCGNTFNCNHPVVRQFIVDCLRYWVTEMHVDGFRFDLASIMTRGSSLWDP 407
          |:| ::|:|| ||| |||:|:||  ::|||||  | |:|||||||||: |    ||
Db    306 TPEGYYFNFSGTGNTLNCNNPIVRGMVLDCLRYWTAEFHIDGFRFDLASILGR-----DP 360

Qy    408 VNVYGAPIEGDMITTGTPLVTPPLIDMISNDPILGGVKLIAEAWDAGGLYQVGQFPHWNV 467
                    | ||  |||:: :: ||||  ||||||||||||||| || :
Db    361 -------------WGYPLANPPLLETLAFDPILARSKLIAEAWDAGGLYQVGSFPSYGR 406

Qy    468 WSEWNGKYRDIVRQFIKGTDGFAGGFAECLCGSPHLYQAGGRKPWHSINFVCAHDGFTLA 527
          |:|||||||| ||:|||  |  |  |: | ||| |||  || |  |||| ||||||||
Db    407 WAEWNGKYRDTVRKFIKGDAGVIGEMAQRLQGSPDLYQGAGRPPSTSINFVTAHDGFTLA 466

Qy    528 DLVTYNKKYNLPNGENNRDGENHNLSWNCGEEGEFARLSVKRLRKRQMRNFFVCLMVSQG 587
          ||| || |:|  |||| || | | ||||| ||      : ||| ||||    |:||||
Db    467 DLVAYNGKHNYANGENGNDGANDNYSWNCGVEGPTDNPDILRLRARQMRNAIALLLVSQG 526

Qy    588 VPMFYMGDEYGHTKGGNNNTYCHDSYVNYFRWDKKEQYSELHRFCCLMTKFR-------- 639
          |||  |||| | |: ||||||||||  |:  |   ||     ||      ||
Db    527 VPMLLMGDEMGKTQDGNNNTYCHDSPFNWLNWHLLEQNQAWFRFVKHCIAFRLAHPVLRN 586

Qy    640 ----KECEGLGLEDFPTAKRLQWHGHQPGKPDWSENSRFVAFSMKDE-------RQGEIY 688
              : |: ||:  ||    : ||| :|   ||| :|| :|| :          :  :||
Db    587 SEHFQNCDYLGV-GFPD---ISWHGVKPWHADWSADSRVLAFMLCGRHAKGGRVKDNQIY 642

Qy    689 VAFNTSHLPAVVELP-ERAGRRWEPVVDTGKPAPYD 723
          || |  :   . |||    | |    :||   | |
Db    643 VAMNMHYESLWFELPAPPVGTTWHVFANTGAQPPED 678
```

```
RESULT   8
Q9M0S5
ID   Q9M0S5      PRELIMINARY;      PRT;   702 AA.
AC   Q9M0S5;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   ISOAMYLASE-LIKE PROTEIN.
GN   AT4G09020.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Spiegel L.A., Huang E.N., Nascimento L.U., de la Bastide M., Vil D.M.,
RA   Preston R.R., Matero A., Shah R., O'Shaughnessy A., Rodriguez M.,
RA   Shekher M., Schutz K., See L.H., Swaby I., Habermann K., Dedhia N.N.,
RA   Mewes H.W., Lemcke K., Mayer K.F.X.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   EU Arabidopsis sequencing project;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL161513; CAB78026.1; -.
DR   HSSP; P10342; 1BF2.
DR   InterPro; IPR000461; Alpha_amylase.
DR   Pfam; PF00128; alpha-amylase; 1.
SQ   SEQUENCE   702 AA;  79558 MW;  264F66C3C30CF942 CRC64;
```

```
  Query Match             36.5%;  Score 1538.5;  DB 10;  Length 702;
```

```
  Best Local Similarity   44.0%;  Pred. No. 2.3e-114;
  Matches  325;  Conservative  110;  Mismatches  222;  Indels   81;  Gaps   19;

Qy      56 RVLAGMPAPLGATALAGGVNFAVYSGGATAAALCLFT---------PEDLKADRVTEEVP 106
           :| :|  :||| : :  |:|||::|  ||:  |||             :|   | : | |
Db       8 KVSSGEVSPLGVSQVDKGINFALFSQNATSVTLCLSLSQRYITSSGKDDTDDDGMIELV- 66

Qy     107 LDPLMNRTGNVWHVFIEG-ELHNMLYGYRFDGTFAPHCGHYLDVSNVVVDPYAKAVISRG 165
           ||| :|:||: ||: :|   |:|:|||||| ||       ||  | | :::|||||| |
Db      67 LDPSVNKTGDTWHICVEDLPLNNVLYGYRVDGPGEWQQGHRFDRSILLLDPYAKLVKGHS 126

Qy     166 EYGVPARGNNCWPQMAGMIPLPYSTFDWEGDLPL-RYPQKDLVIYEMHLRGFTKHDSSNV 224
           :|     :  : |  |      | |||  |      |:||||||::| ||  :|| :
Db     127 SFG---DSSQKFAQFYGTYDFESSPFDWGDDYKFPNIPEKDLVIYEMNVRAFTADESSGM 183

Qy     225 EHP--GTFIGAVSKLDYLKELGVNCIELMPCHEFNELEYSTSSS----KMNFWGYSTINF 278
           :    |:::| : |: :|::||:| :||:|  ||:|||    |:     :| |||||:||
Db     184 DPAIGGSYLGFIEKIPHLQDLGINAVELLPVFEFDELELQRRSNPRDHMVNTWGYSTVNF 243

Qy     279 FSPMTRYTSG---GIKNCGRDAINEFKTFVREAHKRGIE---------VILDVVFNHTAE 326
           |:||:|| ||    ||     |  |||  |:  |  |||          ||||||:||| |
Db     244 FAPMSRYASGEGDPIK-----ASKEFKEMVKALHSAGIEKYSYKFSLQVILDVVYNHTNE 298

Qy     327 GNENGP-ILSFKGVDNTTYYMLAPKGEFYNYSGCGNTFNCNHPVVRQFIVDCLRYWVTEM 385
           ::  |   ||:|:||  |||| |  :  |:|||||| |||||| : |:| ||:|||
Db     299 ADDKYPYTTSFRGIDNKVYYMLDPNNQLLNFSGCGNTLNCNHPVVMELILDSLRHWVTEY 358

Qy     386 HVDGFRFDLASIMTRGSSLWDPVNVYGAPIEGDMITTGTPLVTPPLIDMISNDPILGGVK 445
           ||||||||||:: |                    | |:||  ||||  |: | :|   |
Db     359 HVDGFRFDLASVLCRA------------------TDGSPLSAPPLIRAIAKDSVLSRCK 399

Qy     446 LIAEAWDAGGLYQVGQFPHWNVWSEWNGKYRDIVRQFIKGTDGFAGGFAECLCGSPHLYQ 505
           :||| || |||| ||:||:|: |:|||| ||| ||:|||| .|  | ||  : ||  |||
Db     400 IIAEPWDCGGLYLVGKFPNWDRWAEWNGMYRDDVRRFIKGDSGMKGSFATRVSGSSDLYQ 459

Qy     506 AGGRKPWHSINFVCAHDGFTLADLVTYNKKYNLPNGENNRDGENHNLSWNCGEEGEFARL 565
             |||:| :||| |||||||  |||:|| |:|  |||   || | | |||| |||
Db     460 VNQRKPYHGVNFVIAHDGFTLRDLVSYNFKHNEANGEGGNDGCNDNHSWNCGFEGETGDA 519

Qy     566 SVKRLRKRQMRNFFVCLMVSQGVPMFYMGDEYGHTKGGNNNTYCHDSYVNYFRWDKKEQY 625
           :| || |||:|| : ||:||| ||  ||||||||: ||||:| ||: :| |:| : :
Db     520 HIKSLRTRQMKNFHLALMISQGTPMMLMGDEYGHTRYGNNNSYGHDTSLNNFQWKELDAK 579

Qy     626 SELH-RFCCLMTKFRKECEGLGLEDFPTAKRLQWHGHQPGKPDW-SENSRFVAFSMKDER 683
           : | ||   : |||     |  |:| |   : ||     : :| :  |:|:||:: |
Db     580 KQNHFRFFSEVIKFRHSHHVLKHENFLTQGEITWH-----EDNWDNSESKFLAFTLHDGI 634

Qy     684 QG-EIYVAFNTS--HLPAVVELPERAGRRWEPVVDTGKPAPYDFLTDDLPDRALTIHQFS 740
           | :||||||     : |::  |   |::|  | ||   :| ||: : :   | |
Db     635 GGRDIYVAFNAHDYFVKALIPQPP-PGKQWFRVADTNLESPDDFVREGVAGVADT----- 688

Qy     741 HFLYSNLYPMLSYSSVIL 758
                 | :  :||::|
Db     689 -------YNVAPFSSILL 699
```

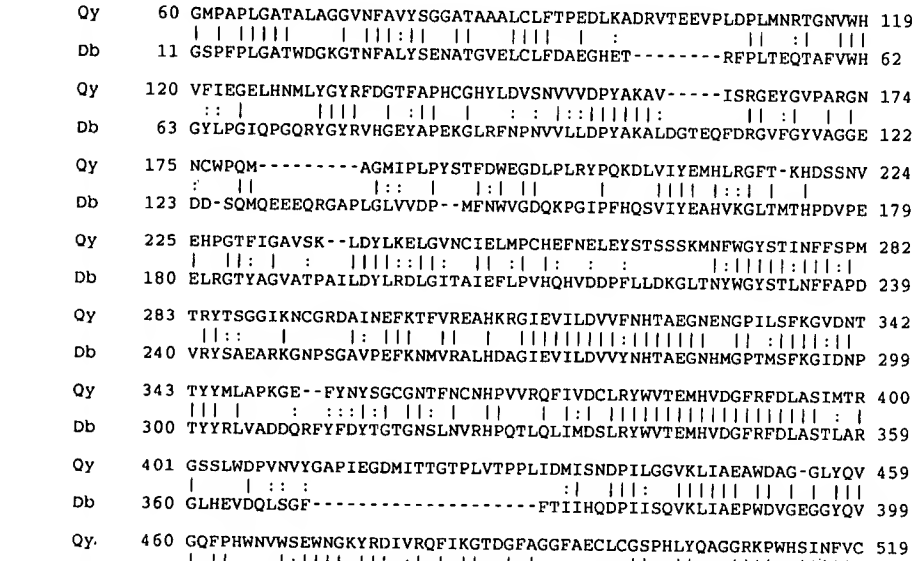
```
RESULT   9
Q9XFG8
ID   Q9XFG8      PRELIMINARY;      PRT;   332 AA.
AC   Q9XFG8;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   ISOAMYLASE 1 (FRAGMENT).
OS   Solanum tuberosum (Potato).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX   NCBI_TaxID=4113;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99364538; PubMed=10437827;
RA   Sun C., Sathish P., Ahlandsberg S., Jansson C.;
RT   "Analyses of isoamylase gene activity in wild-type barley indicate its
RT   involvement in starch synthesis.";
RL   Plant Mol. Biol. 40:431-443(1999).
DR   EMBL; AF142591; AAD33891.1; -.
DR   HSSP; P10342; 1BF2.
DR   InterPro; IPR000461; Alpha_amylase.
DR   Pfam; PF00128; alpha-amylase; 1.
FT   NON_TER       1      1
FT   NON_TER     332    332
SQ   SEQUENCE   332 AA;  37526 MW;  6B63599D1339E751 CRC64;
```

```
  Query Match             35.4%;  Score 1491.5;  DB 10;  Length 332;
  Best Local Similarity   78.9%;  Pred. No. 4.3e-111;
  Matches  262;  Conservative   32;  Mismatches   33;  Indels    5;  Gaps    1;

Qy     196 DLPLRYPQKDLVIYEMHLRGFTKHDSSNVEHPGTFIGAVSKLDYLKELGVNCIELMPCHE 255
           ||||::||:|||||||||:|||| |:||  ::|||::| | |||:|||||||||||||||
Db       1 DLPLKFPQRDLVIYEMHVRGFTNHESSETKYPGTYLGVVEKLDHLKELGVNCIELMPCHE 60

Qy     256 FNELEYST-----SSSKMNFWGYSTINFFSPMTRYTSGGIKNCGRDAINEFKTFVREAHK 310
           |||||| :       || |||||||:|||||| ||:| |: |||  ||||||  |:||||
Db      61 FNELEYYSYNSVLGDSKFNFWGYSTVNFFSPMGRYSSAGLSNCGLGAINEFKYLVKEAHK 120

Qy     311 RGIEVILDVVFNHTAEGNENGPILSFKGVDNTTYYMLAPKGEFYNYSGCGNTFNCNHPVV 370
           ||||||:|||||||||||||||||||:|:||: :| |||||||||||| |||||||:|:|
Db     121 RGIEVIMDVVFNHTAEGNENGPILSFRGIDNSVFYTLAPKGEFYNYSGSGNTFNCNNPMV 180

Qy     371 RQFIVDCLRYWVTEMHVDGFRFDLASIMTRGSSLWDPVNVYGAPIEGDMITTGTPLVTPP 430
           |||||||||||||||| |||||||:|| || |: |||||  |:||:||||||| :||
Db     181 RQFIVDCLRYWVTEMHVDRFRFDLASILTRSSSSWNAVNVYGNSIDGDVITTGTPLTSPP 240

Qy     431 LIDMISNDPILGGVKLIAEAWDAGGLYQVGQFPHWNVWSEWNGKYRDIVRQFIKGTDGFA 490
           |||||||||| |||||||||| |||||||  ||| :||||||||||:||||||||||:
Db     241 LIDMISNDPILRGVKLIAEAWDCGGLYQVGMLPHWGIWSEWNGKYRDMVRQFIKGTDGFS 300

Qy     491 GGFAECLCGSPHLYQAGGRKPWHSINFVCAHD 522
           | |||||||||:||| ||||||:|||||||||
Db     301 GAFAECLCGSPNLYQKGGRKPWNSINFVCAHD 332
```

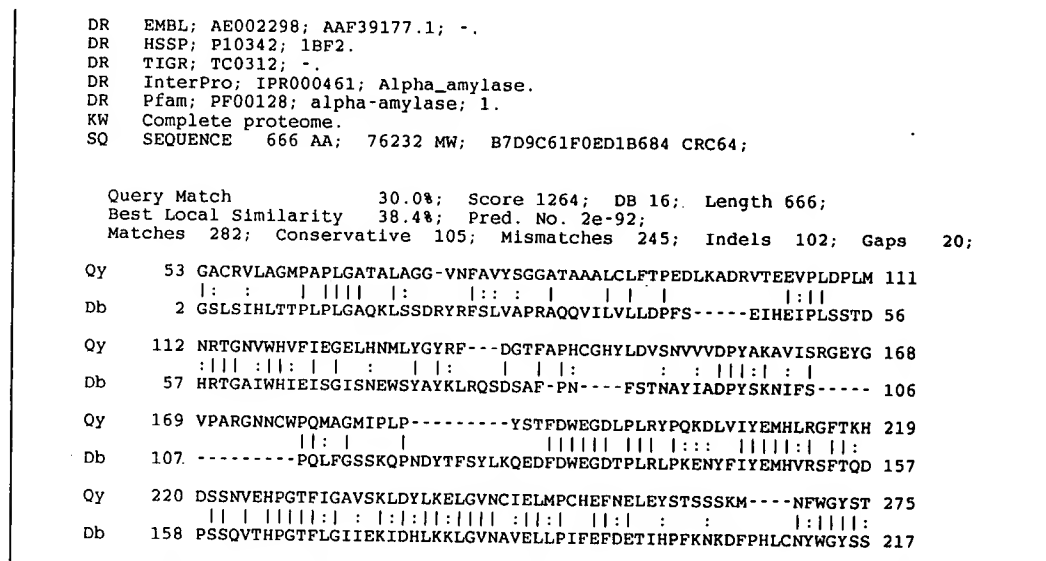
```
RESULT  10
P72691
ID   P72691      PRELIMINARY;      PRT;   746 AA.
AC   P72691;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   GLYCOGEN OPERON PROTEIN GLGX.
GN   GLGX OR SLR0237.
OS   Synechocystis sp. (strain PCC 6803).
OC   Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX   NCBI_TaxID=1148;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97061201; PubMed=8905231;
RA   Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA   Miyajima N., Hirosawa M., Sugiura M., Sasamoto S., Kimura T.,
RA   Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA   Shimpo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA   Tabata S.;
RT   "Sequence analysis of the genome of the unicellular cyanobacterium
RT   Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT   entire genome and assignment of potential protein-coding regions.";
RL   DNA Res. 3:109-136(1996).
DR   EMBL; D90900; BAA16698.1; -.
DR   HSSP; P10342; 1BF2.
DR   InterPro; IPR000461; Alpha_amylase.
DR   InterPro; IPR004193; isoamylase_N.
DR   Pfam; PF00128; alpha-amylase; 1.
DR   Pfam; PF02922; isoamylase_N; 1.
KW   Complete proteome.
SQ   SEQUENCE   746 AA;  83500 MW;  B0416E698B956414 CRC64;
```

```
  Query Match             33.0%;  Score 1388;  DB 16;  Length 746;
  Best Local Similarity   42.0%;  Pred. No. 2.8e-102;
  Matches  305;  Conservative   86;  Mismatches  251;  Indels   84;  Gaps   17;

Qy      60 GMPAPLGATALAGGVNFAVYSGGATAAALCLF-TPEDLKADRVTEEVPLDPLMNRTGNVW 118
           |   |||||  | |||| ::|  |    | ||  | |    |  |   |    |||   |
Db      17 GQTFPLGATVYADGVNFCLFSKHAERVTLLLFDRPNDPAPARTIE---LHRGRNRTFYYW 73

Qy     119 HVFIEGELHNMLYGYRFDGTFAPHCGHYLDVSNVVVDPYAKAVISRGEY---GVPARGNN 175
           |||::|     :| || ||   |  || |  |::|||||:: :  |      | |:|
Db      74 HVFVKGLKAGQVYAYRVDGPHEPEKGHRFDPDKVLLDPYAKAIVGKDIYDRKAAMALGDN 133

Qy     176 CWPQMAGMIPLPYSTFDWEGDLPLRYPQKDLVIYEMHLRGFTKHDSSNVEH--PGTFIGA 233
           |  |    : :  | :||| |    | |    :|||:|: |||:: :| :     ||: |
Db     134 C-AQALRSVVVDTSVYDWEDDHAPRTPYAASIIYELHVGGFTRNPNSGLSENKRGTYAGL 192

Qy     234 VSKLDYLKELGVNCIELMPCHEFNELEYSTSSSKMNFWGYSTINFFSPMTRYTSGGIKNC 293
           : |: |||||:  :||:| | |:  :        |:|||||| ||:|   |::    :
Db     193 IEKIPYLKELGITAVELLPVHYFDPED--AQPGLTNYWGYSTIGFFAPHQGYSAA---DD 247

Qy     294 GRDAINEFKTFVREAHKRGIEVILDVVFNHTAEGNENGPILSFKGVDNTTYYML-APKGE 352
           : ::||:  |:  || |||||||||||||||| || |||:|:|| |||:|   |
Db     248 PLEVVDEFRDMVKALHKAGIEVILDVVFNHTAEGNEKGPTLSFRGIDNRTYYILDEDKSS 307

Qy     353 FYNYSGCGNTFNCNHPVVRQFIVDCLRYWVTEMHVDGFRFDLASIMTRGSSLWDPVNVYG 412
           : |||||||:   |||||   |:| |||||:|||||||||||:: | :          |
Db     308 YSNYSGCGNSVKANHPVVGGLILDSLRYWVSEMHVDGFRFDLASVLVRDTK--------G 359

Qy     413 APIEGDMITTGTPLVTPPLIDMISNDPILGGVKLIAEAWDAGGLYQVGQFPHWNVW-SEW 471
            |: |  | |        :|  | :|||| | ||||||||| ||| ||:|     | :||
Db     360 VPLHGSEIATAN------IIWAIESDPILAGTKLIAEAWDAAGLYSVGKFVELADWFAEW 413

Qy     472 NGKYRDIVRQFIKGTDGFAGGFAECLCGSPHLYQAGGRKPWHSINFVCAHDGFTLADLVT 531
           || :|| ||:|:|| :|     |  | ||| :|       |||||  |||||| |||:
Db     414 NGPFRDDVRRFVKGDNGAVPALASRLLGSPDIYYRQDTDINRSINFVTCHDGFTLVDLVS 473

Qy     532 YNKKYNLPNGENNRDGENHNLSWNCGEEGEFARLSVKRLRKRQMRNFFVCLMVSQGVPMF 591
           ||:|:|  ||| |||| | | ||||| |||     : :|| ||::||   |  ||| ||
Db     474 YNEKHNEANGEKNRDGTNDNFSWNCGVEGETDDPKINQLRLRQIKNFLTILFFSQGTPML 533

Qy     592 YMGDEYGHTKGGNNNTYCHDSYVNYFRWDKKEQYSELHRFC------------------ 632
           |||  | |: |||| || |: :::| |  :| :|:   |
Db     534 LMGDPVGRTQQGNNNGYCQDNELSWFDWSGEETHSDETHFLRGIIALTQSLSLFLEESLL 593

Qy     633 ----CLMTKFRKE---CEGLGLE-DFPTAKRLQ----------------------WHG 660
               : :|  |    || | : : |  | |                        |||
Db     594 PVLNFALPRFPDENDGSEGDGQDGEEPVYKPLPPVDEAALREMYLKATKDLPGSCVVWHG 653

Qy     661 HQPGKPDWSENSRFVAFSMKDERQGEI-YVAFNTSHLPAVVELPE-RAGRRWEPVVDTGK 718
           :  :|||  |  :| ::      || :: ||    |  |||    |  | :|||
Db     654 VKLAQPDWSYTSHTLAVTLFHPGAEEILHLIFNAYWEPLNFELPPLNDGLIWHRLVDTYL 713

Qy     719 PAPYDF 724
           | | ||
Db     714 PTPLDF 719


RESULT  11
Q9Z8F5
ID   Q9Z8F5      PRELIMINARY;      PRT;   664 AA.
AC   Q9Z8F5;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   GLYCOGEN HYDROLASE (DEBRANCHING).
GN   GLGX OR CPN0388 OR CP0367.
OS   Chlamydia pneumoniae (Chlamydophila pneumoniae).
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83558;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CWL029;
RX   MEDLINE=99206606; PubMed=10192388;
RA   Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
RA   Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT   "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RL   Nat. Genet. 21:385-389(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AR39;
RX   MEDLINE=20150255; PubMed=10684935;
RA   Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA   White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA   Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA   Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA   Eisen J., Fraser C.M.;
RT   "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT   pneumoniae AR39.";
RL   Nucleic Acids Res. 28:1397-1406(2000).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=J138;
RX   MEDLINE=20330349; PubMed=10871362;
RA   Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA   Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT   "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT   from Japan and CWL029 from USA.";
RL   Nucleic Acids Res. 28:2311-2314(2000).
DR   EMBL; AE001623; AAD18532.1; -.
DR   EMBL; AE002198; AAF38216.1; -.
DR   EMBL; AP002546; BAA98596.1; -.
DR   HSSP; P10342; 1BF2.
DR   TIGR; CP0367; -.
DR   InterPro; IPR000461; Alpha_amylase.
DR   InterPro; IPR004193; isoamylase_N.
DR   Pfam; PF00128; alpha-amylase; 1.
DR   Pfam; PF02922; isoamylase_N; 1.
KW   Hydrolase; Complete proteome.
SQ   SEQUENCE   664 AA;  75030 MW;  81C0F83E2D32D0D2 CRC64;


  Query Match             31.0%;  Score 1305.5;  DB 16;  Length 664;
  Best Local Similarity   42.6%;  Pred. No. 9.5e-96;
  Matches  279;  Conservative   88;  Mismatches  231;  Indels   57;  Gaps   16;

Qy      61 MPAPLGATALA-GGVNFAVYSGGATAAALCLFTPEDLKADRVTE--EVPLDPLMNRTGNV 117
           :| ||||: ::     ||:|:  ||   | |         |  :|  |||| |  :||| :
Db      10 VPLPLGASKISPNRYRFALYASQATEVILAL-------TDENSEVIEVPLYPDTHRTGAI 62

Qy     118 WHVFIEGELHNMLYGYRFDGTFAPHCGHYLDVSNVVVDPYAKAVISRGEYGVPARGNN-- 175
           ||: |||      | :|  |    | |         : ||||| : |   :|   :  :
Db      63 WHIEIEGISDQSSYAFRVHGP-KKH-GMQYSFKEYLADPYAKNIHSPQSFGSRKKQGDYA 120

Qy     176 -CWPQMAGMIPLPYSTFDWEGDLPLRYPQKDLVIYEMHLRGFTKHDSSNVEHPGTFIGAV 234
            |:        |    | |:|| ||  |::::|||||:| ||:  || | ||||:| :
Db     121 FCY--------LKEEPFPWDGDQPLHLPKEEMIIYEMHVRSFTQSSSSRVHAPGTFLGII 172

Qy     235 SKLDYLKELGVNCIELMPCHEFNELEYSTSSSKM----NFWGYSTINFFSPMTRYTSGGI 290
           |:|:| :||:| :||:|  ||:|  :   :||     |:||: :|||||  ||
Db     173 EKIDHLHKLGINAVELLPIFEFDETAHPFRNSKFPYLCNYWGYAPLNFFSPCRRYAYAS- 231

Qy     291 KNCGRDAINEFKTFVREAHKRGIEVILDVVFNHTAEGNENGPILSFKGVDNTTYYMLAPK 350
             |      |||| |: |: ||||||||||||      |   |  :|  :||:|  :
Db     232 DPCAPS--REFKTLVKTLHQEGIEVILDVVFNHT---GLQGTTCSLPWIDTPSYYILDAQ 286

Qy     351 GEFYNYSGCGNTFNCNHPVVRQFIVDCLRYWVTEMHVDGFRFDLASIMTRGSSLWDPVNV 410
           | | |||||||| | |      |:|:| ||||| ||||||||||||: :|| |
Db     287 GHFTNYSGCGNTLNTNRAPTTQWILDILRYWVEEMHVDGFRFDLASVFSRGPS------- 339

Qy     411 YGAPIEGDMITTGTPLVTPPLIDMISNDPILGGVKLIAEAWDAGGLYQVGQFPHWNV-WS 469
                   |:||   |::: || ||:|   |:||| ||||||||| ||  :  ||
Db     340 -----------GSPLQFAPVLEAISFDPLLASTKIIAEPWDAGGLYQVGYFPTLSPRWS 387

Qy     470 EWNGKYRDIVRQFIKGTDGFAGGFAECLCGSPHLYQAGGRKPWHSINFVCAHDGFTLADL 529
```

```
Db    388 EWNGPYRDNVKAFLNGDQNLIGTFASRISGSQDIYPHG--SPTNSINYVSCHDGFTLCDT 445

Qy    530 VTYNKKYNLPNGENNRDGENHNLSWNCGEEGEFARLSVKRLRKRQMRNFFVCLMVSQGVP 589

Db    446 VTYNHKHNEANGEDNRDGTDANYSYNFGTEGKTEDPGILEVRERQLRNFFLTLMVSQGIP 505

Qy    590 MFYMGDEYGHTKGGNNNTYCHDSYVNYFRWDKKEQYSELHRFCCLMTKFRKECEGLGLED 649

Db    506 MIQSGDEYAHTAEGNNNRWALDSNANYFLWDQLTAKPTLMHFLCDLIAFRKKYKTLFNRG 565

Qy    650 FPTAKRLQWHGHQPGKPDWSENSRFVAFSMKDERQGEIYVAFNTSHLPAVVELPE 704

Db    566 FLSNKEISWVDAMGNPMTWRPGN-FLAFKIKSPK-AHVYVAFHVGAQDQLATLPK 618


RESULT 12
O84046
ID   O84046      PRELIMINARY;      PRT;   666 AA.
AC   O84046;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   GLYCOGEN HYDROLASE (DEBRANCHING).
GN   GLGX OR CT042.
OS   Chlamydia trachomatis.
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX   NCBI_TaxID=813;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=D/UW-3/CX;
RX   MEDLINE=99000809; PubMed=9784136;
RA   Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,
RA   Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,
RA   Davis R.W.;
RT   "Genome sequence of an obligate intracellular pathogen of humans:
RT   Chlamydia trachomatis.";
RL   Science 282:754-759(1998).
DR   EMBL; AE001278; AAC67632.1; -.
DR   HSSP; P10342; 1BF2.
DR   InterPro; IPR000461; Alpha_amylase.
DR   InterPro; IPR004193; isoamylase_N.
DR   Pfam; PF00128; alpha-amylase; 1.
DR   Pfam; PF02922; isoamylase_N; 1.
KW   Hydrolase; Complete proteome.
SQ   SEQUENCE   666 AA;  76076 MW;  E58F03AB064FEC7F CRC64;


  Query Match           30.9%;  Score 1300;  DB 16;  Length 666;
  Best Local Similarity   39.8%;  Pred. No. 2.6e-95;
  Matches  289;  Conservative  102;  Mismatches  241;  Indels   94;  Gaps   20;

Qy     57 VLAGMPAPLGATAL-AGGVNFAVYSGGATAAALCLFTPEDLKADRVTE--EVPLDPLMNR 113

Db      6 VRSTIPLPLGAKKLSADRYRFSLFSSQAQQVTLVLLDP-------LSEIHEIPLSSTDHR 58

Qy    114 TGNVWHVFIEGELHNMLYGYRFDGT------FAPHCGHYLDVSNVVVDPYAKAVISRGEY 167

Db     59 TGAIWHIEIAGISSEWSYAYKLRGTDLSSQKFA--------TDSYIADPYSKNIYSPQLF 110

Qy    168 GVPARGNNCWPQMAGMIPLPYSTFDWEGDLPLRYPQKDLVIYEMHLRGFTKHDSSNVEHP 227

Db    111 GSPKQEKD-----YAFSYLKHEDFDWEGDTPLHLPKENYFIYEMHVRSFTRDPSSQVSHP 165

Qy    228 GTFIGAVSKLDYLKELGVNCIELMPCHEFNELEYSTSSSKM----NFWGYSTINFFSPMT 283

Db    166 GTFLGIIEKIDHLKQLGVHAVELLPIFEFDETVHPFKNQDFPHLCNYWGYSSVNFFCPSR 225

Qy    284 RYTSGGIKNCGRDAINEFKTFVREAHKRGIEVILDVVFNHTAEGNENGPILSFKGVDNTT 343

Db    226 RYTYGA-DPCA--PAREFKTLVKALHRAGIEVILDVVFNHTGFEGTSCPL---PWIDLES 279

Qy    344 YYMLAPKGEFYNYSGCGNTFNCNHPVVRQFIVDCLRYWVTEMHVDGFRFDLASIMTRGSS 403

Db    280 YYMVNDHGDLMNFSGCGNTVNTNTPTTLKWILDALRYWVQEMHVDGFRFDLASVFSR--- 336

Qy    404 LWDPVNVYGAPIEGDMITTGTPLVTPPLIDMISNDPILGGVKLIAEAWDAGGLYQVGQFP 463

Db    337 --DP-------------QGVPLPLTPILQAISSDSILSETKLIAEPWDAGGLYQLGHFP 380

Qy    464 HWNV-WSEWNGKYRDIVRQFIKGTDGFAGGFAECLCGSPHLYQAGGRKPWHSINFVCAHD 522

Db    381 SISTRWSEWNGCYRDHVKAFLNGDAHQVSSFASRISGSHDIYPNG--KPTNSINYICSHD 438

Qy    523 GFTLADLVTYNKKYNLPNGENNRDGENHNLSWNCGEEGEFARLSVKRLRKRQMRNFFVCL 582

Db    439 GFTLYDTVAYNDKHNEENGEYNRDGTSANYSYNFGCEGETTDPTICALRERQMKNFFLAL 498

Qy    583 MVSQGVPMFYMGDEYGHTKGGNNNTYCHDSYVNYFRWDKKEQYSELHRFCCLMTKFRKEC 642

Db    499 FLSQGIPMIQSGDEYGHTAYGNNNHWCLDTKINYFLWDRLAERKELFSFLCQVIALRKAY 558

Qy    643 EGLGLEDFPTAKRLQWHGHQPGKPDWSENSRFVAFSMKDERQGEIYVAFNTSHLPAVVEL 702

Db    559 TELFNTSFLSEDTITWLNTKGSPREWGAD-HYLAFELK-HLNYSLFVAFYSGNERIEISL 616

Qy    703 PE--RAGRRWEPVVD--TGKPAPYDFLTDDLPDRALTIHQFSHFLYSNLYPMLS---YSS 755

Db    617 PKPRKEHLAYEKIVDSTTG----------------------FFSQILSPKLSLEPYSS 652

Qy    756 VILVLR 761

Db    653 LVAISR 658


RESULT 13
Q9RXP5
ID   Q9RXP5      PRELIMINARY;      PRT;   720 AA.
AC   Q9RXP5;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   GLYCOGEN OPERON PROTEIN GLGX.
GN   DR0264.
OS   Deinococcus radiodurans.
OC   Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.
OX   NCBI_TaxID=1299;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=R1;
RX   MEDLINE=20036896; PubMed=10567266;
RA   White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA   Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA   Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA   Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,
RA   Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA   Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA   Fraser C.M.;
RT   "Genome sequence of the radioresistant bacterium Deinococcus
RT   radiodurans R1.";
RL   Science 286:1571-1577(1999).
DR   EMBL; AE001888; AAF09848.1; -.
DR   HSSP; P10342; 1BF2.
DR   TIGR; DR0264; -.
DR   InterPro; IPR000461; Alpha_amylase.
DR   InterPro; IPR004193; isoamylase_N.
DR   Pfam; PF00128; alpha-amylase; 1.
DR   Pfam; PF02922; isoamylase_N; 1.
KW   Complete proteome.
SQ   SEQUENCE   720 AA;  80613 MW;  DF23118CBE3E7157 CRC64;


  Query Match           30.1%;  Score 1268;  DB 16;  Length 720;
  Best Local Similarity   39.8%;  Pred. No. 1.1e-92;
  Matches  282;  Conservative   92;  Mismatches  241;  Indels   94;  Gaps   17;
```



RESULT 14
Q9PKZ6
ID Q9PKZ6 PRELIMINARY; PRT; 666 AA.
AC Q9PKZ6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE GLYCOSYL HYDROLASE FAMILY PROTEIN.
GN TC0312.
OS Chlamydia muridarum.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83560;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MOPN / NIGG;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT pneumoniae AR39.";
RL Nucleic Acids Res. 28:1397-1406(2000).



DR EMBL; AE002298; AAF39177.1; -.
DR HSSP; P10342; 1BF2.
DR TIGR; TC0312; -.
DR InterPro; IPR000461; Alpha_amylase.
DR Pfam; PF00128; alpha-amylase; 1.
KW Complete proteome.
SQ SEQUENCE 666 AA; 76232 MW; B7D9C61F0ED1B684 CRC64;

Query Match 30.0%; Score 1264; DB 16; Length 666;
Best Local Similarity 38.4%; Pred. No. 2e-92;
Matches 282; Conservative 105; Mismatches 245; Indels 102; Gaps 20;

RESULT 15
Q9KY03
ID Q9KY03 PRELIMINARY; PRT; 715 AA.
AC Q9KY03;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

```
DE   PUTATIVE GLYCOGEN DEBRANCHING ENZYME.
GN   GLGX.
OS   Streptomyces coelicolor.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=1902;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2);
RA   Seeger K.J., Harris D.;
RL   Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2);
RA   Thomson N.R., Parkhill J., Barrell B.G., Rajandream M.A.;
RL   Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2);
RX   MEDLINE=97000351; PubMed=8843436;
RA   Redenbach M., Kieser H.M., Denapaite D., Eichner A., Cullum J.,
RA   Kinashi H., Hopwood D.A.;
RT   "A set of ordered cosmids and a detailed genetic and physical map for
RT   the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL   Mol. Microbiol. 21:77-96(1996).
DR   EMBL; AL356932; CAB92884.1; -.
DR   HSSP; P10342; 1BF2.
DR   InterPro; IPR000461; Alpha_amylase.
DR   InterPro; IPR004193; isoamylase_N.
DR   Pfam; PF00128; alpha-amylase; 1.
DR   Pfam; PF02922; isoamylase_N; 1.
SQ   SEQUENCE   715 AA;  80520 MW;  364C09DF6395D2CE CRC64;
```

Query Match 29.7%; Score 1251; DB 2; Length 715;
Best Local Similarity 38.0%; Pred. No. 2.5e-91;
Matches 290; Conservative 92; Mismatches 259; Indels 122; Gaps 20;

```
Qy     57 VLAGMPAPLGATALAGGVNFAVYSGGATAAALCLFTPEDLKADRVTEEVPLDPLMNRTGN 116
Db     13 VWSGHPYPLGAAYDGQGTNFALFSEVAERVDLVL-VDDDGNHSTV-------PLPDVDGF 64

Qy    117 VWHVFIEGELHNMLYGYRFDGTFAPHCGHYLDVSNVVVDPYAKAV---------ISRGEY 167
Db     65 VWHCYLPGVGPGQRYGYRVHGPWAPAVGHRCNPAKLLLDPYTRAVDGLVDNHASLFERAR 124

Qy    168 GVPARGNNCWPQMAGMIPLPYSTFDWEGDLPLRYPQKDLVIYEMHLRGFTK-HDSSNVEH 226
Db    125 GKADPGDSAGHTMLGVVTDPF--FDWGDDRPPRRPYSESVIYEAHVRGLSRTHPDVPEEL 182

Qy    227 PGTFIGAV--SKLDYLKELGVNCIELMPCHEFNELEYSTSSSKMNFWGYSTINFFSPMTR 284
Db    183 RGTYAGLAHPAVVDHLTSLGVTAVELMPVHQFVHDGVLLDRGLSNYWGYNTIGFFAPHNG 242

Qy    285 YTSGGIKNCGRDAINEFKTFVREAHKRGIEVILDVVFNHTAEGNENGPILSFKGVDNTTY 344
Db    243 YAALGTRG---QQVSEFKSMVKTLHEAGLEVILDVVYNHTAEGNERGPTLSFRGIDNASY 299

Qy    345 YMLAPKG--EFYNYSGCGNTFNCNHPVVRQFIVDCLRYWVTEMHVDGFRFDLASIMTRGS 402
Db    300 YRLVDGDWQHYYDTTGTGNSLLMRHPYVLQLIMDSLRYWVTEMHVDGFRFDLAATLARQF 359

Qy    403 SLWDPVNVYGAPIEGDMITTGTPLVTPPLIDMISNDPILGGVKLIAEAWDAG-GLYQVGQ 461
Db    360 HEVDRLSAF------------------FDLIQQDPVISRVKLIAEPWDVGEGGYQVGN 399

Qy    462 FPHWNVWSEWNGKYRDIVRQFIKGTDGFAGGFAECLCGSPHLYQAGGRKPWHSINFVCAH 521
Db    400 FP--QLWSEWNGKYRDAVRDFWRAEDHSLGEFASRLTGSSDLYQHSRRRPRASVNFVTAH 457

Qy    522 DGFTLADLVTYNKKYNLPNGENNRDGENHNLSWNCGEEGEFARLSVKRLRKRQMRNFFVC 581
Db    458 DGFTLRDLVSYNDKHNEANGEDNRDGESHNRSWNCGAEGGTKDPAVRELRGRQQRNFLAT 517

Qy    582 LMVSQGVPMFYMGDEYGHTKGGNNNTYCHDSYVNYFRWDKKEQYSELHRFCCLMTKFRKE 641
Db    518 LLLSQGIPMICHGDELGRTQRGNNNAYCQDNEISWIDWRLDGEQRALLDFARRLIALR-- 575

Qy    642 CEGLGLEDFPTAKRLQW-------HGHQP--------------GKPDWSENSRFV----- 675
Db    576 ------ADHPVLRRRRFFHGETLTHADQPLPDLVWLLPDAREMTDDDWQRSDAHTVGVFL 629

Qy    676 ---AFSMKDER-----QGEIYVAFNTSHLPAVVELPERA-GRRWEPVVDTGKPAPYDFLT 726
Db    630 NGDAIAEPDPRGRPVVDDSFLLLLNSHWEPADFRLPDAGYGERWTALVDTADP------- 682

Qy    727 DDLPDRA-------LTIHQFSHFLYSNLYPMLSYSSVILVLRP 762
Db    683 DGVPDEAEHKAGTRLTVEQ--------------RSLVLLSRP 710
```

Search completed: July 31, 2002, 12:14:40
Job time: 389 sec